CUK-H03252-78-024968

78

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

WEI-LIANG CHOW
THE JOHNS HOPKINS UNIVERSITY

PHILIP HARTMAN
THE JOHNS HOPKINS UNIVERSITY

GILBERT A. HUNT
CORNELL UNIVERSITY

G. D. MOSTOW
YALE UNIVERSITY

WITH THE COÖPERATION OF

R. BOTT
C. CHEVALLEY
K. L. CHUNG
J. A. DIEUDONNÉ
J.-I. IGUSA
A. M. GLEASON
L. V. HÖRMANDER
K. KODAIRA
S. SMALE
A. WEIL

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

510.5
Am 3

Volume LXXXVI, Number 1
JANUARY, 1964

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND 21218
U. S. A.

CONTENTS

PAGE

Notes on homotopy of Thom spectra. By Arunas Liulevicius, . 1

A complete description of the normal subgroups of genus one of the modular group. By Morris Newman, ... 17

On the integral representations of quadratic forms over local fields. By Carl Riehm, ... 25

Quasiconformal mappings and extremal lengths. By H. Renggli, . 63

Generalized automorphic forms and certain holomorphic vector bundles. By Mikio Ise, ... 70

Harmonic mappings of Riemannian manifolds. By James Eells, Jr. and J. H. Sampson, ... 109

Analytic disks in maximal ideal spaces. By John Wermer, . . 161

Spherical functions over 𝔓-adic fields, II. By F. I. Mautner, . . 171

Compactifications of symmetric spaces. By Calvin C. Moore, . . 201

On the graded ring of theta-constants. By Jun-ichi Igusa, . . 219

The American Journal of Mathematics appears four times yearly.

The subscription price of the Journal is $11.00 in the U. S.; $11.30 in Canada and $11.60 in other foreign countries. The price of single numbers is $3.00.

Manuscripts intended for publication in the Journal should be sent to Professor W. L. Chow, The Johns Hopkins University, Baltimore, Md. 21218.

Subscriptions to the Journal and all business communications should be sent to The Johns Hopkins Press, Baltimore, Maryland 21218.

Claims for replacement of missing numbers must be received within 3 months (6 months in the case of foreign subscribers) following publication of the number.

The Johns Hopkins Press supplies to the authors 100 free reprints of every article appearing in the American Journal of Mathematics. On the other hand, neither The Johns Hopkins Press nor the American Journal of Mathematics can accept orders for additional reprints. Authors interested in securing more than 100 reprints are advised to make arrangements directly with the printers, J. H. Furst Co., 109 Market Place, Baltimore, Maryland 21202.

The typescripts submitted can be in English, French, German or Italian and should be prepared in accordance with the instructions listed on the inside back cover of this issue.

Second-class postage paid at Baltimore, Maryland.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST CO., BALTIMORE, MARYLAND

AMERICAN
JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

WEI-LIANG CHOW
THE JOHNS HOPKINS UNIVERSITY

PHILIP HARTMAN
THE JOHNS HOPKINS UNIVERSITY

GILBERT A. HUNT
CORNELL UNIVERSITY

G. D. MOSTOW
YALE UNIVERSITY

WITH THE COÖPERATION OF

R. BOTT
C. CHEVALLEY
K. L. CHUNG
J. A. DIEUDONNÉ
J.-I. IGUSA
A. M. GLEASON
L. V. HÖRMANDER
K. KODAIRA
S. SMALE
A. WEIL

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

PUBLISHED UNDER THE JOINT AUSPICES OF
THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LXXXVI

1964

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND 21218
U. S. A.

INDEX

PAGE

Ahlfors, Lars V. Finitely generated Kleinian groups, ... 413
Ash, M. E. The basic estimate of the $\bar{\partial}$-Neumann problem in the non-Kahlerian case, ... 247
Ax, James. Zeros of polynomials over finite fields, ... 255
Browder, Felix E. Strongly non-linear parabolic biundary value problems, ... 339
Browder, Felix E. Non-local elliptic boundary value problems, . 735
Burgess, C. E. Properties of certain types of wild surfaces in E^3, . 325
Chow, Wei-Liang. On unmixedness theorem, ... 799
Eells, James, Jr. and J. H. Sampson. Harmonic mappings of Riemannian manifolds, ... 109
Ernest, John. A new group algebra for locally compact groups, . 467
Frink, Orrin. Compactifications and semi-normal spaces, . . . 602
Hagis, Peter, Jr. Partitions into odd and unequal parts, . . . 317
Harish-Chandra. Invariant distributions on Lie algebras, . . . 271
Harish-Chandra. Invariant differential operators and distributions on a semisimple Lie algebra, ... 534
Hartman, Philip. Geodesic parallel coordinates in the large, . . 705
Helgason, S. Fundamental solutions of invariant differential operators on symmetric spaces, ... 565
Hochschild, G. and G. D. Mostow. On the algebra of representation functions of an analytic group, II, ... 869
Hsiang, W. C. and R. H. Szczarba. On the tangent bundle of a Grassman manifold, ... 698
Igusa, Jun-ichi. On the graded ring of theta-constants, ... 219
Igusa, Jun-ichi. On Siegel modular forms of genus two (II), . 392
Ise, Mikio. Generalized automorphic forms and certain holomorphic vector bundles, ... 70
Jones, B. Frank. A class of singular integrals, ... 441
Knopp, Marvin Isadore. On generalized Abelian integrals of the second kind and modular forms of dimension zero, . . 430
Kodaira, K. On the structure of compact complex analytic surfaces, I, 751
Kuranishi, Masatake. Sheaves defined by differential equations and application to deformation theory of pseudo-group structures, 379
Kwun, Kyung Whan and Frank Raymond. Almost acyclic maps of manifolds, ... 638
Lang, Serge. Diophantine approximations on toruses, ... 521
Liulevicius, Arunas. Notes on homotopy of Thom spectra, . . 1
Mackie, A. G. On Riemann's method and a variation by Martin, . 728

PAGE

MATSUSAKA, T. and D. MUMFORD. Two fundamental theorems on deformations of palarized varieties, ... 668

MAUTNER, F. I. Spherical functions over $\mathfrak{P}$-adic fields, ... 171

MITCHELL, BARRY. The full imbedding theorem, ... 619

MOORE, CALVIN C. Compactifications of symmetric spaces, . . . 201

MOORE, CALVIN C. Compactifications of symmetric spaces II the Cartan domains, ... 358

MOSTOW, G. D. See G. Hochschild and G. D. Mostow, page 869.

MUMFORD, D. See T. Masustaka and D. Mumford, page 668.

NEWMAN, D. J. The closure of translates in l^p, ... 651

NEWMAN, MORRIS. A complete description of the normal subgroups of genus one of the modular group, ... 17

PUTNAM, C. R. Commutators, absolutely continuous spectra, and singular integral operators, ... 310

RAYMOND, FRANK. See Kyung Whan Kwun and Frank Raymond, page 638.

RENGGLI, H. Quasiconformal mappings and extremal lengths, . . . 63

RIEHM, CARL. On the integral representations of quadratic forms over local fields, ... 25

SAH, CHIH-HAN. A note on Hermitian forms over fields of characteristic 2, ... 262

SAMPSON, J. H. See James Eells, Jr. and J. H. Sampson, page 109.

STEIN, KARL. Maximale holomorphe und meromorphe abbildungen, II, 823

SZCZARBA, R. H. On tangent bundles of fibre spaces and quotient spaces, ... 685

SZCZARBA, R. H. See W. C. Hsiang and R. H. Szczarba, page 698.

WEND, DAVID V. V. Branched regular curve families and finite asymptotic values, ... 493

WERMER, JOHN. Analytic disks in maximal ideal spaces, . . . 161

WINTER, DAVID L. Finite groups having a faithful representation of degree less than $(2p+1)/3$, ... 608

NOTES ON HOMOTOPY OF THOM SPECTRA.

By Arunas Liulevicius.*

Let G be one of the classical Lie groups: O, U, SO, SU, Sp. Let MG be the corresponding Thom spectrum [8]. We wish to compute the homotopy of MG modulo elements of finite order prime to 2. As in [8], the main tool is the Adams spectral sequence [1] for the prime two.

Most of the results of Chapter II were known to Milnor [8] and Novikov [11]. The only claim to originality is Chapter II.

The first chapter is devoted to homological algebra: a "change of rings" theorem is presented, which computes the cohomology of sundry subalgebras of the Steenrod algebra. The second chapter applies these results to compute the integral and 2-primary parts of the homotopy of MG.

It is shown that all differentials vanish in the Adams spectral sequence for MO, MU, MSO. It is also shown that not all differentials are zero in the spectral sequence for MSU and MSp.

Chapter I.

1.1. A change of rings theorem. Let R, S be rings with unit, h: $R \to S$ a ring homomorphism. If N is a left S-module, we may consider it as a left R-module through the homomorphism h. Take (C, d_o, ϵ_o) to be an S-projective resolution of N:

$$(1.1.1) \qquad 0 \leftarrow N \xleftarrow{\epsilon_0} C_0 \xleftarrow{d_0} C_1 \leftarrow \cdots \leftarrow C_q \xleftarrow{d_0} \cdots.$$

Let (X_q, d_0, ϵ_q) be an R-projective resolution of C_q in (1.1.1):

$$(1.1.2) \qquad 0 \leftarrow C_q \xleftarrow{\epsilon_q} X_{q,0} \xleftarrow{d_0} X_{q,1} \leftarrow \cdots \leftarrow X_{q,r} \xleftarrow{d_0} .$$

There exist maps [3] $d_1 \colon X_{q+1} \to X_q$ such that

$$(1.1.3) \qquad \epsilon_q d_1 = d_0 \epsilon_{q+1}, \qquad d_1 d_0 + d_0 d_1 = 0.$$

Received May 21, 1962.
Revised September 16, 1963.

* The author holds a National Science Foundation post-doctoral fellowship.

Let $Y = \sum_q X_q$, $Y_s = \sum_{q+r=s} X_{q,r}$. Define $\epsilon\colon Y \to N$ by $\epsilon = \epsilon_0(\sum_q \epsilon_q)$.

THEOREM I. 1. *There is a differential d in Y such that (Y, d, ϵ) is an R-projective resolution of N.*

Proof. The differential d in Y need only have the following property: if $d = \sum_{k=0}^{\infty} d_k$, where $d_k\colon X_q \to X_{q-k}$, then it is only necessary that d_0 and d_1 are as in (1. 1. 3). The fact that d is a differential is expressed by

$$(1.1.4) \qquad \sum_{i=0}^{k} d_{k-i} d_i = 0, \text{ for all } k.$$

The mapping $d_k\colon X_{q,r} \to X_{q-k,r+k-1}$ is constructed by induction on r. We suppose d_i for $i < k$ have been defined on $X_{q,r}$ and d_k is defined for $X_{q,r-1}$ (note that $d_k = 0$ if $r - 1 < 0$). Let $f_k = -\sum_{i=0}^{k-1} d_{k-i} d_i$. Then it is immediate from (1. 1. 4) that $d_0 f_k = 0$, thus, since X_{q-k} is acyclic, there is an R-map d_k such that $d_0 d_k = f_k$, which is the relation (1. 1. 4) for d_k at $X_{q,r}$.

To show that Y is acyclic under d, we filter Y by $F_p Y = \sum_{j \leqq p} X_j$. In the spectral sequence for this filtration $E_p{}^0 = X_p$, and the differential corresponds to d_0. Thus $E_p{}^1 = C_p$, and the differential d^1 corresponds to d_0 in virtue of (1. 1. 3). Thus $E_p{}^2 = E_p = 0$ if $p \neq 0$, and $E_0 = N$ in total degree 0.

Remark. Theorem I. 1 is stated by Wall in [14] for the special case $R = Z(G)$, $S = Z(H)$, where G, H are finite groups and H is a quotient of G. The proof above is an immediate generalization.

We will be interested in Theorem I. 1 applied to Hopf algebras. Let E be a graded, connected, associative Hopf algebra [9] over a field K; let F be a Hopf subalgebra of E.

THEOREM I. 2 (Milnor-Moore). *E is free as a left (or right) F-module.*

COROLLARY I. 3. *$E \otimes_F$ is an exact functor.*

This is Theorem 2. 5 of [9]. The corollary is immediate.

We say that F is normal [9] in E (and write $F \lhd E$) if $E\bar{F} = \bar{F}E$, where $\bar{F}$ denotes the augmentation ideal of F. If $F \lhd E$, then $B = E/E\bar{F}$ is again a connected Hopf algebra over K (B is denoted by $E/\!/F$). Let $\pi\colon E \to B$ be the projection.

PROPOSITION I. 4. *If (D, d, ϵ) is an F-free resolution of K, then $(E \otimes_F D,\ 1 \otimes_F d,\ 1 \otimes_F \epsilon)$ is an E-free resolution of B.*

Proof. As a left E-module, $B = E \otimes_F K$. The complex $E \otimes_F D$ is E-free and is still acyclic, according to Corollary I. 3.

COROLLARY I. 5. *If E, B, F are as above*

$$\operatorname{Ext}_B{}^{s,t}(B,K) \cong \operatorname{Ext}_B{}^{s,t}(K,K). \tag{1.1.5}$$

Proof. Let D be as in Proposition I. 4. Then $\operatorname{Hom}_B(E \otimes_F D, K) \cong \operatorname{Hom}_F(D,K)$ as cochain complexes, hence (1.1.5).

Let us apply Theorem I. 1 to $R = E$, $S = B$, F as above. Suppose that N is a graded B-module. We let $C_i = B \otimes_K \bar{C}_i$ as left B-modules (projective = free, since B is graded, connected [4]). According to Proposition I. 4 we can take for X_q the complex $E \otimes_K \bar{C}_q \otimes_K \bar{D}$, where $\bar{D} = K \otimes_F D$. We let $X_{i,j} = E \otimes_K \bar{C}_i \otimes_K \bar{D}_j$. To construct the differential d, it is sufficient to give E-maps d_k satisfying (1.1.3) and (1.1.4).

We need just one more remark before we go on to examine some subalgebras of the Steenrod algebra. Let E be a connected Hopf algebra over the field K, F a Hopf subalgebra. Let J be a left ideal in F, $I = EJ$ the E-extension of J to a left ideal in E.

LEMMA I. 6. *$E \otimes_F F/J = E/I$ as a left E-module.*

Proof. Consider the exact sequence of F-modules

$$0 \to J \to F \to F/J \to 0$$

and apply $E \otimes_F$.

COROLLARY I. 7.

$$\operatorname{Ext}_E{}^{s,t}(E/I,K) \cong \operatorname{Ext}_F{}^{s,t}(F/J,K).$$

Proof. Take an F-free resolution of F/J and apply $E \otimes_F$, then use Lemma I. 6 and Corollary I. 3.

1. 2. Some subalgebras of the Steenrod algebra. Let A be the Steenrod algebra over Z_2. According to Milnor [7], the graded dual A^* is a polynomial algebra on generators ξ_n, $\operatorname{grade}(\xi_n) = 2^n - 1$, $n = 1, 2, \cdots$. If I is a sequence of non-negative integers, Sq^I will denote the element in the Milnor basis [7] corresponding to I. In particular, if we let Δ_i be the sequence which consists of 1 in the i-th place, and zeroes elsewhere, we define $Q_i = \mathrm{Sq}^{\Delta_i}$, $R_i = \mathrm{Sq}^{2\Delta_i}$, $\operatorname{grade}(Q_i) = 2^{i+1} - 1$, $\operatorname{grade}(R_i) = 2^{i+2} - 2$. It is easy to verify that the elements Q_i, R_j satisfy the following relations (here $[x,y] = xy + yx$):

$$\begin{aligned} &[Q_i, Q_j] = 0, \\ &Q_i Q_i = 0, \\ &[R_j, R_j] = 0 \quad (i, j \neq 0), \\ &[R_0, R_j] = Q_0 Q_{j+1} \quad (j \neq 0), \\ &R_0 R_0 = Q_0 Q_1, \\ &[Q_i, R_j] = 0 \quad (i \neq 0), \\ &[Q_0, R_j] = Q_{j+1}. \end{aligned} \tag{1.2.1}$$

Let C be the subalgebra of A generated by 1 and Q_j, $j = 0, 1, \cdots$, B the subalgebra of A generated by 1, Q_0, $R_0, \cdots, R_j, \cdots$ for $j = 0, 1, \cdots$.

Let $\alpha^* \colon A^* \to A^*$ be the squaring map in the dual of the Steenrod algebra—it is a Hopf algebra monomorphism. It is clear that $A^* /\!/ \alpha^*(A^*)$ is an exterior algebra on the classes of the generators ξ_n. The graded dual of $A^* /\!/ \alpha^*(A^*)$ can be identified with C. Consider the dual of α^*:

$$\alpha \colon A \to A. \tag{1.2.2}$$

According to Lemma 2.4.2 of [2] we have proved statement (i) of the following proposition:

PROPOSITION I. 8. *Let A, B, C, as above. Then $C \subset B \subset A$ are inclusions of Hopf algebras, and*

(i) $C \lhd A$, Kernel $\alpha = A\bar{C}$,
(ii) $B \lhd A$, Kernel $\alpha \cdot \alpha = A\bar{B}$.

The proof of (ii) is just as easy, and is left to the reader.

1. 3. Computing Ext. We will need the groups $\mathrm{Ext}_E(Z_2, Z_2)$, where E is B, C, A_0, A_1, $C^{\#}$, where B, C were defined in 1. 2, A_n is the subalgebra of A generated by 1 and Sq^{2^j}, $j \leqq n$, and $C^{\#}$ is the subalgebra generated by A_1 and C.

The algebra $\mathrm{Ext}_C^{*,*}(Z_2, Z_2)$ is well known [7]: it is a polynomial algebra in generators $q_i \in \mathrm{Ext}_C^{1, 2^{i+1}-1}(Z_2, Z_2)$. We shall use the standard minimal resolution [8] of Z_2 over C. Generators will be in one-to-one correspondence with finitely nonzero sequences of integers I (the free C-generator corresponding to I will be denoted by $[I]$). Let $I = (i_0, i_1, \cdots, i_k, \cdots)$; then degree $[I] = \sum_k i_k$, grade $[I] = \sum_k . i_k(\mathrm{B}^{k+1} - 1)$. The differential in the minimal resolution is defined by

$$d[I] = \sum_{r=0}^{\infty} Q_r [I - \Delta_r], \tag{1.3.1}$$

where we set $[I - \Delta_r] = 0$ if $i_r = 0$.

Let us compute $\operatorname{Ext}_B(Z_2, Z_2)$ using Theorem I. 1. Since $C \subset B$, $\operatorname{Ker} \alpha \mid B = B\bar{C}$ and $C \lhd B$. We have that $\alpha(B) = C'$, where C' is C with the grading of every element multiplied by two. For the module $X_{i,j}$ in (1. 1. 2) we take (see remarks after Corollary I. 5) the free B-module on generators $[I] \otimes [J]$, where degree $[I] = i$, degree $[J] = j$, and grade $[I] \otimes [J] = 2\,\text{grade}[I] + \text{grade}[J]$. The augmentation ϵ_i (1. 1. 2) is defined by $\epsilon_i([I] \otimes [J]) = 0$ if degree $[J] \neq 0$, $\epsilon_i([I] \otimes [0]) = [I]$. Both d_0 and d_o are defined by (1. 3. 1).

Proposition I. 9. *We can define the maps d_k for $k \geqq 1$ as follows:*

$$d_1([I] \otimes [J]) = \sum_k R_k[I - \Delta_k] \otimes [J]$$

$$+ \sum_k (j_{k+1} + 1)[I - \Delta_k] \otimes [J - \Delta_0 + \Delta_{k+1}],$$

$$d_2[I] \otimes [J]) = \sum_i (j_{i+1} + 1) Q_0[I - \Delta_0 - \Delta_i] \otimes [J + \Delta_{i+1}],$$

$$d_3([I] \otimes [J])$$

$$= \sum_{k<t} (j_{k+1} + 1)(j_{t+1} + 1)[I - \Delta_0 - \Delta_k - \Delta_t] \otimes [J + \Delta_{k+1} + \Delta_{t+1}]$$

$$+ \sum_t C_2{}^{j_{t+1}+2}[I - \Delta_0 - 2\Delta_t] \otimes [J + 2\Delta_{t+1}],$$

$$d_n = 0 \text{ for } n \geqq 4;$$

where j_k denotes the k-th entry in J, and $C_s{}^r$ is the binomial coefficient $r!/s!(r-s)!$

Proof. We have to show (see Theorem I. 1) that

$$\epsilon_Q d_1 = d_o \epsilon_{Q+1},$$

$$\sum_{i=0}^{k} d_{k-i} d_i = 0, \qquad k = 1, 2, \cdots.$$

The first is easy to verify: $\epsilon([I] \otimes [J]) = 0$, unless $J = 0$, and $\epsilon(\vartheta[I] \otimes [0]) = \alpha(\vartheta)[I]$, thus

$$\epsilon d_1([I] \otimes [J]) = \sum_k Q_k \epsilon([I - \Delta_k] \otimes [J] = d_o \epsilon([I] \otimes [J]).$$

The second set of equations is more laborious to verify:

(a) to show $d_0 d_1 + d_1 d_0 = 0$;

$$d_1 d_0([I] \otimes [J]) = d_1(\sum_s Q_s[I] \otimes [J - \Delta_s])$$

$$= \sum_{s,k} Q_s R_k[I - \Delta_k] \otimes [J - \Delta_s]$$

$$+\sum_{s\neq k+1}(j_{k+1}+1)Q_s[I-\Delta_k]\otimes[J-\Delta_0-\Delta_s+\Delta_{k+1}]$$

$$+\sum_k j_{k+1}Q_{k+1}[I-\Delta_k]\otimes[J-\Delta_0].$$

According to (1.2.1) $Q_sR_k=R_kQ_s$ if $s\neq 0$, $Q_0R_k=R_kQ_0+Q_{k+1}$, thus

$$d_1d_0([I]\otimes[J]=\sum_{s,k}R_kQ_s[I-\Delta_k]\otimes[J-\Delta_s]$$

$$+\sum_{s,k}(j_{k+1}+1)Q_s[I-\Delta_k]\otimes[J-\Delta_0-\Delta_s+\Delta_{k+1}]$$

$$=d_0d_1([I]\otimes[J]),$$

which proves (a).

(b) To show $d_0d_2+d_1d_1+d_2d_0=0$, we examine each term separately:

$$d_0d_2([I]\otimes[J])=\sum_{i,s}(j_{i+1}+1)Q_sQ_0[I-\Delta_0-\Delta_i]\otimes[J-\Delta_s+\Delta_{i+1}],$$

$$d_1d_1([I]\otimes[J])=\sum_{i,s}R_sR_i[I-\Delta_s-\Delta_i]\otimes[J]$$

$$=\sum_{s<i}[R_s,R_i][I-\Delta_s-\Delta_i]\otimes[J]$$

$$+\sum_i R_iR_i[I-2\Delta_i]\otimes[J],$$

$$=\sum_i Q_{i+1}Q_i[I-\Delta_0-\Delta_i]\otimes[J],$$

$$d_2d_0([I]\otimes[J])=d_2(\sum_s Q_s[I)\otimes[J-\Delta_s])$$

$$=\sum_{s\neq i+1}(j_{i+1}+1)Q_sQ_0[I-\Delta_0-\Delta_i]\otimes[J-\Delta_s+\Delta_{i+1}]$$

$$+\sum_i j_{i+1}Q_{i+1}Q_0[I-\Delta_0-\Delta_i]\otimes[J].$$

Thus (b) is verified.

(c) To show $d_0d_3+d_1d_2+d_2d_1+d_3d_0=0$:

$$d_0d_3([I]\otimes[J])=\sum_s\sum_{k<t}(j_{k+1}+1)(j_{t+1}+1)Q_s[I-\Delta_0-\Delta_k-\Delta_t]$$

$$\otimes[J-\Delta_s+\Delta_{k+1}+\Delta_{t+1}]$$

$$+\sum_{s,t}C_2{}^{j_{t+1}+2}Q_s[I-\Delta_0-2\Delta_t]\otimes[J+2\Delta_{t+1}-\Delta_s],$$

$$d_1d_2([I]\otimes[J])=d_1(\sum_t(j_{t+1}+1)Q_0[I-\Delta_0-\Delta_t]\otimes[J+\Delta_{t+1}])$$

$$=\sum_{t,k}(j_{t+1}+1)Q_0R_k[I-\Delta_0-\Delta_t-\Delta_k]\otimes[J+\Delta_{t+1}]$$

$$+\sum_{k\neq t}(j_{k+1}+1)(j_{t+1}+1)Q_0[I-\Delta_0-\Delta_t-\Delta_k]\otimes[J+\Delta_{t+1}+\Delta_{k+1}-\Delta_0],$$

$$\begin{aligned} d_2 d_1([I] \otimes [J]) &= d_2\Big(\sum_k R_k[I-\Delta_k] \otimes [J] + \sum_k (j_{k+1}+1)[I-\Delta_k] \\ &\qquad\qquad \otimes [J-\Delta_0+\Delta_{k+1}]\Big) \\ &= \sum_{k,t} (j_{t+1}+1) R_k Q_0 [I-\Delta_0-\Delta_t-\Delta_k] \otimes [J+\Delta_{t+1}] \\ &\quad + \sum_{k \neq t} (j_{k+1}+1)(j_{t+1}+1) Q_0 [I-\Delta_0-\Delta_t-\Delta_k] \otimes [J-\Delta_{t+1}-\Delta_0+\Delta_{k+1}]. \end{aligned}$$

Thus

$$(d_1 d_2 + d_2 d_1)([I] \otimes [J]) = \sum_{t,k} (j_{t+1}+1) Q_{k+1}[I-\Delta_0-\Delta_t-\Delta_k] \otimes [J+\Delta_{t+1}].$$

$$\begin{aligned} d_3 d_0([I] \otimes [J]) &= d_3\Big(\sum_s Q_s[I] \otimes [J-\Delta_s]\Big) \\ &= \sum_{\substack{s-1 \neq k,t \\ k<t}} (j_{k+1}+1)(j_{t+1}+1) Q_s [I-\Delta_0-\Delta_k-\Delta_t] \otimes [J-\Delta_s+\Delta_{k+1}+\Delta_{t+1}] \\ &\quad + \sum_{k<t} (j_{k+1})(j_{t+1}+1) Q_{k+1}[I-\Delta_0-\Delta_k-\Delta_t] \otimes [J+\Delta_{t+1}] \\ &\quad + \sum_{k<t} (j_{k+1}+1)(j_{t+1}) Q_{t+1}[I-\Delta_0-\Delta_k-\Delta_t] \otimes [J+\Delta_{k+1}] \\ &\quad + \sum_{s \neq t+1} C_2{}^{j_{t+1}+2} Q_s [I-\Delta_0-2\Delta_t] \otimes [J+2\Delta_{t+1}-\Delta_s] \\ &\quad + \sum_t C_2{}^{j_{t+1}+1} Q_{t+1}[I-\Delta_0-2\Delta_t] \otimes [J+\Delta_{t+1}]. \end{aligned}$$

Thus (c) is verified, for $C_2{}^{j+1} + (j+1) = C_2{}^{j+2}$.

(d) We ask the reader to verify that

$$\begin{aligned} d_1 d_3 + d_3 d_1 &= 0, \\ d_2 d_2 &= 0, \\ d_2 d_3 + d_3 d_2 &= 0, \\ d_3 d_3 &= 0, \end{aligned}$$

which completes the proof of Proposition I. 9.

$\mathrm{Ext}_B(Z_2, Z_2)$ is now computed by taking the cohomology of $\mathrm{Hom}_B(Y, Z_2)$, which can be identified with a polynomial algebra

$$P = Z_2[x_0, x_1, \cdots, x_i, \cdots, y_1, \cdots, y_j, \cdots], \tag{1.3.2}$$

where x_i is the element in the dual basis to the $[I] \otimes [J]$ corresponding to $[\Delta_i] \otimes [0]$, and y_j corresponds to the element $[0] \otimes [\Delta_j]$; thus degree x_i $=$ degree $y_j = 1$. grade $x_i = 2 \cdot$ grade $y_i = 2^{i+2} - 2$. Let $\delta_i = \mathrm{Hom}_B(d_i, 1)$. Then acocrding to Proposition I. 7 we have $\delta_i = 0$ for $i \neq 1, 3$; δ_1 is a derivation, $\delta_1(x_i) = \delta_1(y_0) = 0$, $\delta_1(y_j) = y_0 x_{j-1}$. Furthermore, δ_3 is a map of

$Z_2[x_0, x_1, \cdots, y\ ]$-modules, $\delta_3(x_i) = \delta_3(y_0) = 0$; if J is a sequence of finitely non-zero non-negative integers, we denote by y^J the monomial in the y_k's, where the exponent of y_k is j_k. Then

$$\delta_3(y^J) = \sum_{i<t} j_i j_t x_0 x_{i-1} x_{t-1} y^{(J-\Delta_i-\Delta_t)}$$
$$+ \sum_t C_2{}^{j_t} x_0 x_{t-1}{}^2 y^{(J-2\Delta_t)}.$$

We also have $\delta_1\delta_1 = 0$, $\delta_1\delta_3 + \delta_3\delta_1 = 0$, $\delta_3\delta_3 = 0$, (for (1.1.4) holds).

LEMMA I.10. *Under the differential* δ_1, $P = M \otimes_{Z_2} N$, *where*

$$M = Z_2[y_1{}^2, y_2{}^2, \cdots, y_j{}^2, \cdots]$$

with zero differential, and

$$N = Z_2[x_0, x_1, \cdots, x_i, \cdots, y_0] \otimes_{Z_2} \Lambda[y_1, y_2, \cdots, y_j, \cdots]$$

with differential δ *such that* δ *is a derivation in* N, $\delta(x_j) = \delta(y_0) = 0$, $\delta(y_j) = y_0 x_{j-1}$. (Here $\Lambda[a_j]$ denotes the exterior algebra over Z_2 on elements a_j.)

Proof. Immediate by inspection.

The homology of N is easily computed. Since multiplication by y_0 is a monomorphism in N, all δ-cycles will be found by examining the δ'-cycles in the complex $N/(y_0)$, where δ' is the derivation defined by $\delta'(x_i) = 0$, $\delta'(y_j) = x_{j-1}$. This, however, is precisely the Koszul complex ([3], p. 153), which is known to be acyclic—it is precisely a minimal resolution of Z_2 over the augmented ring $R = Z_2[x_0, x_1, \cdots, x_i, \cdots]$. In particular, the only δ'-cycles are R-multiples of the basic cycles

$$(1.3.3) \qquad \delta'(y^J) = \sum_k x_{j_k-1} y^{(J-\Delta_k)},$$

where now J is a sequence of zeroes (first entry 0) and a finite number of ones. The only R-relations among the cycles $\delta'(y^J)$ are R-linear combinations of the relations giving $\delta'\delta' = 0$. Return now to the complex N. Let x_i, y_0 denote the classes in $H(N, \delta_1)$ determined by x_i, y_0, respectively. Let $U(J)$ be the class of $\delta'(y^J)$ (1.3.3). We have the following relations in $H(N, \delta_1)$:

$$(1.3.4) \qquad \begin{aligned} &y_0 x_i = 0 \qquad i = 0, 1, \cdots, \\ &\sum_i x_{i-1} U(J - \Delta_i) = 0, \\ &y_0 U(J) = 0, \end{aligned}$$

all other relations are R-linear combinations of these. We thus have:

PROPOSITION I. 11. *The homology of* P (1. 3. 2) *under* δ_1 *is a* $Z_2[x_0, x_1, \cdots, x_i, \cdots, y_0]$*-module with*

$$Z_2[y_1^2, \cdots, y_j^2, \cdots] \otimes_{Z_2} \Lambda[U(J)]$$

as generating subspace. The basic relations over $Z_2[x_0, \cdots, x_i, \cdots, y_0]$ *are given in* (1. 3. 4).

The homology of $H(P, \delta_1)$ under the differential induced by δ_3 is harder to handle. The method used will be illustrated in two special cases.

Let B_k be the subalgebra of B generated by 1, Q_0 and $R_0, \cdots, R_k$. A resolution for Z_2 over B_k is obtained by taking the resolution in Proposition I. 9 and setting $[I] \otimes [J] = 0$ if some $i_r \neq 0$ for $r > k$ or some $j_s \neq 0$ for $s > k+1$. In the cochain complex P (1. 3. 2) this corresponds to setting $y_j = 0$ if $j > k+1$, $x_i = 0$ if $i > k$. It is clear from the given form of the differential $\delta_1 + \delta_3$ in P that

$$(1.3.5) \qquad i^*\colon \mathrm{Ext}_B{}^{s,t}(Z_2, Z_2) \to \mathrm{Ext}_{B_k}{}^{s,t}(Z_2, Z_2)$$

is an epimorphism, and for $t - s < 2^{k+2} - 3$ it is an isomorphism, because $B = B_k$ in gradings $< 2^{k+2} - 2$.

For our purposes it will be sufficient to compute $\mathrm{Ext}_{B_1}(Z_2, Z_2)$. We have to determine the homology of

$$(1.3.6) \qquad P_1 = Z_2[x_0, x_1, y_0, y_1, y_2]$$

under the differential $\delta_1 + \delta_3$, where δ_1 and δ_3 are as in (1. 3. 2). According to Proposition I. 11, $H(P_1, \delta_1)$ is a $Z_2[x_0, x_1, y_0]$-module with $Z_2[y_1^2, y_2^2] \otimes \Lambda[U(0,1,1)]$ as a generating set and $x_0 y_0 \cdot 1 = 0$, $x_1 y_0 \cdot 1 = 0$, $y_0 U(0,1,1) = 0$ as basic relations.

We can decompose $H(P_1, \delta_1)$ under δ_3 into the following form:

$$(1.3.7) \qquad K \otimes [L \otimes_T M + Q + N],$$

where $K = Z_2[y_1^4, y_2^4]$, $d_K = 0$; $T = Z_2[x_0]$, $L = T \otimes \Lambda[x_1] \otimes \Lambda[U(0,1,1)]$, $d_L = 0$; $M = Z_2[x_0, x_1^2] \otimes \Lambda[y_1^2, y_2^2]$, d_M derivation, $d_M(x_i) = 0$, $d_M(y_j^2) = x_0 x_{j-1}^2$; $Q = Z_2[y_0]$, $d_Q = 0$; $N = \bar{Q} \otimes \Lambda[y_1^2, y_2^2]$, $d_N = 0$. The homology of M is readily computable: it is a $Z_2[x_0, x_1^2]$-module with generators 1 and α, where $\alpha = \{x_0^2 y_2^2 + x_1^2 y_1^2\}$, and relations $x_0 \alpha = 0$, $x_0 x_1^2 \cdot 1 = 0$, $x_0^3 \cdot 1 = 0$.

Since T is a principal ideal ring, we can readily compute $H(L \otimes_T M)$, by using the Künneth theorem [3]. Since L is T-free, and $H(L) = L$, we have $H(L \otimes_T M) = H(L) \otimes_T H(M)$.

Let us define the following classes in $\mathrm{Ext}_{B_1}(Z_2, Z_2)$:

$$q_0 = \{y_0\}, \qquad k_0 = \{x_0\}, \qquad k_1 = \{x_1\},$$

$$\tau_0 = \{y_0 y_1^2 + x_0^2 y_1\}, \qquad \tau_1 = \{y_0 y_2^2 + x_0 x_1 y_2\},$$

$$\tau_{01} = \{y_0 y_1^2 y_2^2 + x_0^2 y_1 y_2^2 + x_0 x_1 y_1^2 y_2\},$$

$$\gamma = \{x_0 y_2 + x_1 y_1\}, \qquad \omega_1 = \{y_1^4\}, \qquad \omega_2 = \{y_2^4\}.$$

PROPOSITION I. 12. *As an algebra,* $\mathrm{Ext}_{B_1}(Z_2, Z_2)$ *is free as a* $Z_2[\omega_1, \omega_2]$*-module on generators*

$$\begin{gathered} q_0^n,\ k_1^{n+1},\ q_0^n\tau_0,\ q_0^n\tau_1,\ q_0^n\tau_{01}, \\ k_1^n\gamma,\ k_1^n\gamma^2,\ k_1^n\gamma^3, \quad n = 0, 1, 2, \cdots \\ k_0,\ k_0^2,\ k_0k_1,\ k_0k_1^2,\ k_0\gamma,\ k_0^2\gamma,\ k_0k_1\gamma,\ k_0k_1^2\gamma; \end{gathered} \tag{1.3.8}$$

furthermore, the elements satisfy the following (incomplete) set of relations

$$\begin{gathered} \tau_0^2 = q_0^2\omega_1, \quad k_0^3 = 0, \quad k_0^2k_1 = 0, \quad k_0\tau_0 = 0, \\ k_0\gamma^2 = 0, \quad q_0k_0 = 0, \quad q_0k_1 = 0, \quad q_0\gamma = k_0^2k_1. \end{gathered} \tag{1.3.9}$$

Proof. Our computation $H(P_1, \delta_1; \delta_3)$ proves Proposition I. 12 additively. It remains to check that the appropriate classes represent the products claimed in Ext. This is easy to see from the resolution given in Proposition I. 7 and the Yoneda definition of products (see [2]).

COROLLARY I. 13. $\mathrm{Ext}_{B_0}(Z_2, Z_2)$ *is a free* $Z_2[\omega_1]$*-module with the following set of classes as free generators*:

$$q_0^n,\ q_0^n\tau_0,\ k_0,\ k_0^2; \tag{1.3.10}$$

furthermore, all relations are consequences of the following

$$k_0\tau_0 = 0,\ k_0^3 = 0,\ q_0k_0 = 0,\ \tau_0^2 = q_0^2\omega_1. \tag{1.3.11}$$

Proof. Trivial computation.

Since $B_0 = A_0$, we have computed $\mathrm{Ext}_{A_0}(Z_2, Z_2)$. It is now easy to compute $\mathrm{Ext}_{C^\#}(Z_2, Z_2)$. Let H be the subalgebra of $C^\#$ generated by 1 and Q_j, $j = 2, 3, \cdots$. According to relations (1.2.1), the elements of H and A_0 commute in $C^\#$, and

$$C^\# = A_0 \otimes H \tag{1.3.12}$$

as algebras. The algebra $\mathrm{Ext}_H(Z_2, Z_2)$ is isomorphic to $Z_2[q_2, \cdots, q_j, \cdots]$ $j = 2, 3, \cdots$, degree$(q_j) = 1$, grade$(q_j) = 2^{j+1} - 1$. Since [3] (1.3.12) implies that

$$\mathrm{Ext}_{C^\#}(Z_2, Z_2) \cong \mathrm{Ext}_{A_0}(Z_2, Z_2) \otimes \mathrm{Ext}_H(Z_2, Z_2) \tag{1.3.13}$$

as algebras, we have from Corollary I.13:

COROLLARY I.14. $\mathrm{Ext}_{C^\#}(Z_2, Z_2)$ *is a free* $Z_2[\omega_1, q_2, q_3, \cdots]$*-module with the set of classes* (1.3.10) *as free generators and* (1.3.11) *as a complete set of relations.*

Chapter II.

2.1. Cohomology of Thom spectra. Let $\boldsymbol{MO}$ be the Thom spectrum [8] for the orthogonal group; similarly, let $\boldsymbol{MU}$, $\boldsymbol{MSp}$, $\boldsymbol{MSO}$, $\boldsymbol{MSU}$ be the Thom spectra for the unitary, symplectic, special unitary groups, respectively. Let

$$\begin{aligned} &L = H^*(\boldsymbol{MO}; Z_2), && M = H^*(\boldsymbol{MU}; Z_2), \\ &N = H^*(\boldsymbol{MSp}; Z_2), && U = H^*(\boldsymbol{MSO}; Z_2), \\ &&V = H^*(\boldsymbol{MSU}; Z_2). \end{aligned} \tag{2.1.1}$$

Let A be the Z_2 Steenrod algebra and $\alpha: A \to A$ the dual of the squaring map in A^* (1..2).

PROPOSITION II.1 (Novikov). *There exist* Z_2*-module isomorphisms* $f: M \to L$, $g: N \to M$, $h: V \to U$ *such that the following diagrams are commutative*:

$$\begin{tikzcd} A \otimes N \arrow[r] \arrow[d, "\alpha \otimes f"'] & N \arrow[d, "g"] \\ A \otimes M \arrow[r] \arrow[d, "\alpha \otimes g"'] & M \arrow[d, "f"] \\ A \otimes L \arrow[r] & L \end{tikzcd} \qquad \begin{tikzcd} A \otimes V \arrow[r] \arrow[d, "\alpha \otimes h"'] & V \arrow[d, "h"] \\ A \otimes U \arrow[r] & U \end{tikzcd}$$

where the horizontal arrows indicate the action of A.

Proof. We see at once that it suffices to prove the proposition for $L = H^*(RP^\infty; Z_2)$, $M = H^*(CP^\infty; Z_2)$ $N = H^*(HP^\infty; Z_2)$, where RP^∞, CP^∞, HP^∞ indicate the infinite-dimensional projective spaces over the reals, coplexes and quaternions, respectively, with f and g algebra isomorphisms defined by $f(y) = x$, $g(u) = y$, where x, y, u are the generators of the polynomial algebras. But this is easy, using [7]: if $\vartheta \in A$,

$$\vartheta x = \sum_n \langle \xi_n, \vartheta \rangle x^{2^n},$$

and $h\colon M \to L$, $k\colon N \to M$ defined by $h(y) = x^2$, $k(u) = y^2$ are algebra monomorphisms consistent with the action of A. But [7]:

$$\vartheta x^2 = \sum_n \langle \xi_n{}^2, \vartheta \rangle x^{2^{n+1}},$$

hence, since h is an A-homomorphism,

$$\vartheta y = \sum_n \langle \xi_n, \alpha(\vartheta) \rangle y^{2^n},$$

or $f(\vartheta y) = \alpha(\vartheta) f(y)$. Similarly $g(\vartheta u) = \alpha(\vartheta) g(u)$, and the proposition is proved.

Remarks. 1. The above proposition is proved in [11]. 2. L, M, N, U, V carry natural co-algebra structures (see [10]). The maps f, g, h in Proposition II. 1, above, preserve the co-algebra structures.

THEOREM II. 2 (Thom). *As an A-module and coalgebra,* $L = A \otimes S$, *where* $S^* = Z_2[x_2, \cdots, x_k, \cdots]$ $k \neq 2^t - 1$ *for any t.*

Proof. That L is a free A-module was proved in [12]. A new proof which exhibits the co-algebra structure is given in [6]. The proof that L is A-free is very simple. We just notice (this was pointed out by W. Browder) that the theorem of Milnor and Moore (Theorem I. 2, above) holds if E is just a coalgebra, since the proof in [9] only uses the diagonal in E.

THEOREM II. 3 (Wall). *As an A-module, U is a direct sum of a free A-module and of modules with a single generator u and one basic relation* $\mathrm{Sq}^1 u = 0$. *The generators u are in one-to-one correspondence with partitions of non-zero multiples of four into multiples of four.*

This theorem is proved in [13] in an indirect way from the structure of the unoriented cobordism ring $\mathfrak{N}_*$. A direct proof is given in [15].

If H is a graded vector space, let us denote by H' the vector space isomorphic to H but with every grading multiplied by two. Let $H'' = (H')'$. Consider A' (resp. A'') as a left A-module via the map α (respectively $\alpha \circ \alpha$).

THEOREM II. 4 (Milnor, Novikov). *As left A-modules*

i) $M = A' \otimes S'$,

ii) $N = A'' \otimes S''$,

iii) *V is a direct sum of A-modules isomorphic to* A' *and modules isomorphic to* $A'/A'\mathrm{Sq}^1$; *the latter occur in dimensions which are in one-to-one correspondence with non-zero partitions of* 8 *into multiples of* 8.

Proof. Proposition II.1 and Theorems II.2, II.3. We remark that as an A-module $A'/A'\mathrm{Sq}^1$ can be described as follows: it has a single generator u, and for $\vartheta \in A$, $\vartheta u = 0$ if and only if $\vartheta \in \operatorname{Ker} \alpha + A\,\mathrm{Sq}^2 = A\bar{C}^{\#}$.

In order to compute $\operatorname{Ext}_A(H, Z_2)$, where H is one of the modules L, M, N, U, V, it therefore suffices to compute Ext for the A-modules A, A', A'', $A/A\mathrm{Sq}^1$, $A'/A'\mathrm{Sq}^1$, respectively.

PROPOSITION II. 5.

i) $\operatorname{Ext}_A(A, Z_2) = 1$, $1 \in \operatorname{Ext}^{0,0}$;

ii) $\operatorname{Ext}_A(A', Z_2) = Z_2[q_0, \cdots, q_k, \cdots]$, $q_k \in \operatorname{Ext}^{1, 2^{k+1}-1}$;

iii) $\operatorname{Ext}_A^{s,t}(A'', Z_2)$ *is given by Proposition* I.12 *for* $t - s < 13$;

iv) $\operatorname{Ext}_A^{s,t}(A/A\mathrm{Sq}^1, Z_2) = Z_2[q_0]$, $q_0 \in \operatorname{Ext}^{1,1}$;

v) $\operatorname{Ext}_A(A'/A'\mathrm{Sq}^1, Z_2)$ *is given by Corollary* I.14.

Proof. i) is obvious; ii) follows from Corollary I.5 and Proposition I.8; iii) from Corollary I.5, Proposition I.8 and (1.3.5); iv) and v) from Corollary I.7 applied to A and A_0, $C^{\#}$, respectively.

2.2. The Adams spectral sequence. Let X and Y be finite CW complexes with base points. Let $S^m X = X \# S^m$ be the m-th suspension. Define the stable track group $\{X, Y\}_n$ to be $\operatorname{Dir\,lim}[S^{m+n}X, S^m Y]$ under suspensions. Let p be a prime; denote by $H^*(X)$ the augmented cohomology with Z_p coefficients. Let K_n be the subgroup of $\{X, Y\}_n$ consisting of the elements whose order is finite and prime to p.

THEOREM II. 6 (Adams). *There exists a spectral sequence* $\{E_r^{s,t}, d_r\}$, $E_r^{s,t} = E_r^{s,t}(X, Y; p)$ *such that*

$$E_2^{s,t} = \operatorname{Ext}_A^{s,t}(H^*(Y), H^*(X)),$$

$$d_r\colon E_r^{s,t} \to E_r^{s+r, t+r-1},$$

$$\{X, Y\}_n = B^{0,n} \supset B^{1,n+1} \supset \cdots \supset B^{s,n+s} \supset \cdots,$$

where $\bigcap_s B^{s,n+s} = K_n$, $B^{s,t}/B^{s+1,t+1} = E_\infty^{s,t}$, *and* $E_\infty^{s,t}$ *is the limit of* $E_r^{s,t}$ *as* $r \to \infty$.

The theorem is proved in [1]. The theorem remains true if Y is replaced by a spectrum satisfying certain finiteness conditions (see [8]). We shall consider the case $X = S^0$, $Y = \boldsymbol{MG}$, where $\boldsymbol{G}$ is $\boldsymbol{O}$, $\boldsymbol{U}$, $\boldsymbol{Sp}$, $\boldsymbol{SO}$, $\boldsymbol{SU}$, respectively; then $\{S^0, \boldsymbol{MG}\}_n = \operatorname{Dir\,lim} \pi_{n+k}(MG_k)$.

If we let $X = S^0$, $Y = S^0$, we obtain the spectral sequence for the stable homotopy groups of spheres. The Adams spectral sequence for the sphere operates on the spectral sequence for $X = S^0$, Y a spectrum. This can be described as follows. Let H be an A-module, then $\mathrm{Ext}_A(H, Z_p)$ is a right $\mathrm{Ext}_A(Z_p, Z_p)$-module: that is, there is a natural map

$$\mu\colon \mathrm{Ext}_A^{u,v}(H, Z_p) \otimes_{Z_p} \mathrm{Ext}_A^{s,t}(Z_p, Z_p) \to \mathrm{Ext}_A^{u+s,v+t}(H, Z_p)$$

which gives the module structure. We write $\alpha * \beta$ for $\mu(\alpha \otimes \beta)$. We shall need the following:

THEOREM II. 7 (Adams). *Let* $\alpha \in \mathrm{Ext}_A^{u,v}(H^*(Y), Z_p)$, $\beta \in \mathrm{Ext}_A^{s,t}(Z_p, Z_p)$. *Suppose* $d_k(\beta) = 0$ *in the spectral sequence for the sphere for* $k = 2, 3, \cdots$. *Suppose* $d_j(\alpha) = 0$ *for* $j = 2, \cdots, r-1$. *Then*

$$d_r\{\alpha * \beta\} = \{d_r\alpha\} * \beta. \tag{2.2.1}$$

For a much more general result see the proof of Theorem 2. 2 in 1.

Let $p = 2$. Let $q_0 \in \mathrm{Ext}_A^{1,1}(Z_2, Z_2)$ be defined by Sq^1. It is well known that $q_0^n \neq 0$ for all n, and that q_0 is a d_k-cycle for all k in the spectral sequence for the sphere. Furthermore, $* q_0$ corresponds to multiplication by 2 in the E_∞ term.

2. 3. The spectral sequences for the Thom spectra. Let $p = 2$.

THEOREM II. 8. *All differentials in the Adams spectral sequence for* ***MO, MU, MSO*** *are identically zero.*

Proof. According to Proposition II. 5. i) and Theorem II. 2, $E_2^{s,t} = 0$ if $s > 0$ for ***MO***; hence all differentials are zero for a trivial reason.

According to Theorem II. 4. i) and Proposition II. 5. ii), $E_2^{s,t} = 0$ if $t - s$ is odd in the case of ***MU***. Hence again all differentials are zero, since d_k changes the parity of $t - s$.

The case of ***MSO*** is equally easy. The modules $A/A\mathrm{Sq}^1$ give rise to "towers" obtained by multiplying with q_0 (according to Proposition II. 5. iv)). The towers are never in adjacent values of $t - s$, according to Theorem II. 3. It remains to show that if $\alpha \in E_r^{s,t}$, $\beta \in E_r^{s-r,t-r1}$, $\alpha * q_0 \neq 0$, $\beta * q_0 = 0$, then $d_r(\beta) \neq \alpha$. This, however, is immediate from Theorem II. 7: $d_r(\beta) = \alpha$ implies $d_r(\beta * q_0) = \alpha * q_0$, hence $\alpha * q_0 = 0$, which is a contradiction.

THEOREM II. 9. *Not all differentials in the Adams spectral sequence for* ***MSU*** and ***MSp*** are identically zero.

Proof. In both cases $E_2^{0,8} = Z_2$. If we can show that the image of the stable mod 2 Hurewicz homomorphism is zero in dimension 8, then some differential must be non-zero on $E_r^{0,8}$. The assertion about the Hurewicz homomorphism is proved by considering the characteristic numbers of eight-dimensional manifolds which admit a stable ***Sp***- or ***SU***-structure on their tangent bundles. In particular, such manifolds are stably complex, hence their Chern numbers satisfy the following congruences [5]:

$$-2c_4+c_3c_1\equiv 0 \mod 4,$$
$$-c_4+c_3c_1+3c_2^2+4c_2c_1^2-c_1^4\equiv 0 \mod 720,$$
$$c_2c_1^2+2c_1^4\equiv 0 \mod 12.$$

For a stable ***SU***-manifold $c_1=0$, and since we are only interested in the characteristic numbers of a stable ***Sp***-manifold, we may as well assume that it is 3-connected (remark of R. Lashof), hence also here all characteristic numbers involving c_1 are zero. The congruences above therefore give

$$-2c_4\equiv 0 \mod 4,$$
$$-c_4+3c_2^2\equiv 0 \mod 720;$$

in particular,

$$c_4\equiv 0 \mod 2,$$
$$c_2^2\equiv 0 \mod 2,$$

and in both cases the Hurewicz homomorphism in dimension 8 is zero mod 2, which completes the proof of the theorem.

Let us consider $\mathrm{Ext}_A^{s,t}(A'',Z_2)$ for $t-s<9$. Using Proposition II. 5.iii) we exhibit these groups in the following table.

TABLE

8									⋮
7									⋮
6					⋮				$\omega_1 q_0^2$
5	⋮				$\tau_0 q_0^2$				$\omega_1 q_0$
4	⋮				$\tau_0 q_0$				ω_1
3	q_0^3				τ_0			$k_0^2 k_1$	$k_0\gamma$
2	q_0^2		k_0^2				$k_0 k_1$	γ	
1	q_0	k_0				k_1			
0	1								
$s\uparrow$									
$t-s\rightarrow$	0	1	2	3	4	5	6	7	8

It is clear that all differentials for $t - s < 7$ vanish (arguments as for Theorem II. 8). It is easy to determine the group extensions in the E term. According to 8 there is no p-torsion for $p > 2$ in the homotopy of **MSp**, thus we have:

THEOREM II. 10. *The first seven homotopy groups of* **MSp** *are*:

k	0	1	2	3	4	5	6
π_k	Z	Z_2	Z_2	0	Z	Z_2	Z_2

THE UNIVERSITY OF CHICAGO
AND
THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] J. F. Adams, "On the structure and applications of the Steenrod algebra," *Commentarii Mathematici Helvetici*, vol. 32 (1958), pp. 180-214.

[2] ———, "On the non-existence of elements of Hopf invariant one," *Annals of Mathematics*, vol. 72 (1960), pp. 20-104.

[3] H. Cartan and S. Eilenberg, *Homological Algeba*, Princeton University Press, 1956.

[4] S. Eilenberg, "Homological dimension and syzygies," *Annals of Mathematics*, vol. 64 (1956), pp. 328-336.

[5] F. Hirzebruch, "Komplexe Mannigfaltigkeiten," *Proceedings of the International Congress of Mathematicians* 1958 (Edinburgh), pp. 119-136.

[6] A. Liulevicius, "A proof of Thom's theorem," *Commentarii Mathematici Helvetici*, vol. 37 (1962), pp. 121-131.

[7] J. Milnor, "The Steenrod algebra and its dual," *Annals of Mathematics*, vol. 67 (1958), pp. 150-171.

[8] ———, "On the cobordism ring Ω_* and a complex analogue," *American Journal of Mathematics*, vol. 82 (1960), pp. 505-521.

[9] J. Milnor and J. C. Moore, "On the structure of Hopf algebras" (to appear).

[10] S. P. Novikov, "Some problems in the topology of manifolds connected with the theory of Thom spaces," *Doklady Akademii Nauk SSSR*, vol. 132 (1960), pp. 1031-1034 (translation in *Soviet Mathematics*, vol. 1 (1961), pp. 717-719).

[11] ———, "Homotopy properties of Thom complexes" (dissertation; unpublished).

[12] R. Thom, "Quelques proprietes globales des varieties differentiables," *Commentarii Mathematici Helvetici*, col. 28 (1954), pp. 17-86.

[13] C. T. C. Wall, "Determination of the cobordism ring," *Annals of Mathematics*, vol. 72 (1960), pp. 292-311.

[14] ———, "Resolutions for extensions of groups," *Proceedings of the Cambridge Philosophical Society*, vol. 57 (1961), pp. 251-255.

[15] ———, "A characterization of simple modules over the Steenrod algebra mod 2," *Topology*, vol. 1 (1962), pp. 249-254.

A COMPLETE DESCRIPTION OF THE NORMAL SUBGROUPS OF GENUS ONE OF THE MODULAR GROUP.*

By MORRIS NEWMAN.

Let Γ be the 2×2 modular group: that is, the totality of linear fractional transformations

$$\tau' = \frac{a\tau + b}{c\tau + d}$$

where a, b, c, d are integers and $ad - bc = 1$. It is known that Γ is the free product of a cyclic group of order 2 and a cyclic group of order 3,

$$\Gamma = \{x\} * \{y\}, \quad x^2 = y^3 = 1.$$

Γ may be represented as the multiplicative group of 2×2 rational integral matrices of determinant 1 in which a matrix and its negative are identified. If u is an element of Γ then $\bar{u}$ will denote the matrix corresponding to u under this representation. Then the representation may be chosen so that

$$\bar{x} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \bar{y} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$$

and setting

$$z = xy,$$

$$\bar{z} = \begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

We shall write $\{u, v, \cdots\}$ for the group generated by $u, v, \cdots$. Thus $\Gamma = \{x, y\}$.

Let G be a normal subgroup of Γ of finite index μ, and define $n = n(G)$ as the least positive integer n such that the substitution

$$\tau' = \tau + n$$

(equivalently z^n) is in G. Then except for the groups Γ, Γ^2, Γ^3 (Γ^r is the subgroup of Γ generated by the r-th powers of the elements of Γ) the genus g

$$g = 1 + \mu \frac{n - 6}{12n}$$

Received December 11, 1962.

* The preparation of this paper was supported by the Office of Naval Research.

of the associated standard fundamental region of G is given by the formula (see [1] for an excellent discussion of this topic). Thus if $g \neq 1$, it follows that there are only finitely many possibilities for μ and n and consequently only finitely many possibilities for G, since it is known that a finitely generated group possesses only finitely many subgroups of a given finite index (see [2]). For $g = 1$ however no condition is imposed on μ, but n must be 6; and there are actually infinitely many normal subgroups of Γ of finite index and genus 1. These can be classified by uniformization theorems, but the results tell almost nothing about their group-theoretic structure. It will be the purpose of this article to determine their group-theoretic structure completely in terms of three integer parameters.

The first and second commutator subgroups of Γ will be denoted by Γ', Γ'' respectively. Then (see [4] and [6])

Lemma 1. *The commutator subgroup Γ' of Γ is a free group of rank 2, and*

$$(\Gamma : \Gamma') = 6, \Gamma = \sum_{r=0}^{5} z^r \Gamma', \Gamma' = \{xyxy^2, xy^2xy\}. \tag{1}$$

We set

$$a = xyxy^2, \qquad b = xy^2xy.$$

Then a and b have the matrix representations

$$\tilde{a} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \tilde{b} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

From Lemma 1 it follows that Γ'' is a free group of countably infinite rank and of infinite index in Γ' (see [3], vol. 2). Furthermore, Γ'/Γ'' is the free abelian group of rank 2 on the generators $a\Gamma''$, $b\Gamma''$.

If w is any word of Γ' then $e_a(w)$, $e_b(w)$ will denote the sum of the exponents of a, b respectively. Γ'' is characterized by $e_a(w) = e_b(w) = 0$. If u, v are elements of Γ then we shall write

$$u \equiv v \text{ if and only if } uv^{-1} \in \Gamma''.$$

Thus if $u, v \in \Gamma'$ then $u \equiv v$ if and only if

$$e_a(u) = e_a(v), \qquad e_b(u) = e_b(v).$$

We will need

Lemma 2. *Let $\Delta(6)$ be the smallest normal subgroup of Γ containing z^6. Then $\Delta(6) = \Gamma''$.*

The proof is given in [5]. It depends on the observation that Γ' is generated by a and b, and the commutator $ab^{-1}a^{-1}b \in \Delta(6)$. Thus $\Delta(6) \supset \Gamma''$. But $\Delta(6)$ is contained in any normal subgroup of Γ containing z^6 and $z^6 \in \Gamma''$, so that $\Delta(6) \subset \Gamma''$. Hence the conclusion.

LEMMA 3. *Let w be any word of Γ and define the conjugacy*

$$S(w) = zwz^{-1}.$$

Then S is determined from

$$S(a) = ab^{-1}, \qquad S(b) = a. \tag{2}$$

The proof is straightforward, and follows from the decomposition of Γ modulo Γ' given in (1), Lemma 1, and a small computation.

LEMMA 4. *Let G be a normal subgroup of Γ'. Then G is a normal subgroup of Γ if and only if $zGz^{-1} \subset G$.*

The proof follows directly from the decomposition of Γ modulo Γ' given in (1), Lemma 1.

Lemma 4 is used in the proofs of the theorems that follow to determine under what condtions a normal subgroup of Γ' is also a normal subgroup of Γ, occasionally without explicit mention.

LEMMA 5. *Let G be a normal subgroup of Γ, $G \supset \Gamma''$. Suppose that $z^{r_1}a^{s_1}b^{t_1} \in G$, $z^{r_2}a^{s_2}b^{t_2} \in G$. Then $z^{r_1+r_2}a^{s_1+s_2}b^{t_1+t_2} \in G$.*

Proof. Since G is a normal subgroup of Γ and $z^{r_2}a^{s_2}b^{t_2} \in G$ it follows that $a^{s_2}b^{t_2}z^{r_2} \in G$. Hence $z^{r_1}a^{s_1}b^{t_1}a^{s_2}b^{t_2}z^{r_2} \in G$, and thus $z^{r_1+r_2}a^{s_1}b^{t_1}a^{s_2}b^{t_2} \in G$. But $a^{s_1}b^{t_1}a^{s_2}b^{t_2} \equiv a^{s_1+s_2}b^{t_1+t_2}$, and $\Gamma'' \subset G$.
Thus the lemma follows.

We are now ready to prove our first theorem.

THEOREM 1. *Let G be a normal subgroup of Γ of genus 1. Then*

$$\Gamma' \supset G \supset \Gamma''.$$

Proof. Since G is a normal subgroup of Γ of genus 1, the exponent of z modulo G is 6. Thus $z^6 \in G$ and it follows that $G \supset \Delta(6)$. Hence Lemma 2 implies that $G \supset \Gamma''$.

Now suppose that g is any element of G. By Lemma 1 we may write

$$g = z^r g', g' \in \Gamma', \quad 0 \leqq r \leqq 5 \tag{3}$$

and hence

$$g = z^r a^s b^t g'', \quad g'' \in \Gamma''.$$

Since $\Gamma'' \subset G$, it follows that

$$w = z^r a^s b^t \in G.$$

Since G is a normal subgroup of $\Gamma, S(w) \in G$. Using Lemma 3 and keeping in mind that $G \supset \Gamma''$ we obtain

$$\begin{aligned} w &\equiv z^r a^s b^t \in G, \\ S(w) &\equiv z^r a^{s+t} b^{-s} \in G, \\ S^2(w) &\equiv z^r a^t b^{-s-t} \in G, \\ S^3(w) &\equiv z^r a^{-s} b^{-t} \in G, \\ S^4(w) &\equiv z^r a^{-s-t} b^s \in G, \\ S^5(w) &\equiv z^r a^{-t} b^{s+t} \in G. \end{aligned}$$

Then Lemma 5 implies that $z^{2r} \in G$ (since w and $S^3(w)$ belong to G) and that $z^{3r} \in G$ (since $S(w)$, $S^3(w)$ and $S^5(w)$ belong to G). It follows that $z^r \in G$. But z is of exponent 6 modulo G and $0 \leqq r \leqq 5$. Hence $r = 0$, and (3) implies that $g \in \Gamma'$. Hence $G \subset \Gamma'$ and the proof the theorem is complete.

The determination of the normal subgroups of Γ of finite index and of genus 1 is therefore reduced to the determination of the normal subgroups of Γ of finite index contained between Γ' and Γ''. Throughout the remainder of this article we shall assume that G is a normal subgroup of Γ (not necessarily of finite index) such that $\Gamma' \supset G \supset \Gamma''$, $G \neq \Gamma''$. We are going to study the quotient group G/Γ'' which is abelian since $\Gamma' \supset G \supset \Gamma''$.

We first prove

THEOREM 2. *G/Γ'' is the free abelian group on 2 generators.*

Proof. Γ'/Γ'' is the free abelian group on the two generators $a\Gamma''$, $b\Gamma''$ and G/Γ'' is a subgroup of Γ'/Γ''. Hence G/Γ'' is either cyclic or the free abelian group on two generators. Assume that G/Γ'' is cyclic, with generator $a^s b^t \Gamma''$.

Since G is a normal subgroup of Γ there must be an integer k such that

$$\begin{aligned} S(a^s b^t) &\equiv (a^s b^t)^k, \\ (ab^{-1})^s a^t &\equiv (a^s b^t)^k. \end{aligned} \tag{4}$$

From (4) we obtain that

$$\begin{aligned} (k-1)s - t &= 0, \\ s + kt &= 0, \end{aligned}$$

which is possible if and only if $s = t = 0$. This implies that $G = \Gamma''$, which is excluded. Thus G/Γ'' cannot be cyclic. Hence G/Γ'' is the free abelian group on two generators $A\Gamma''$, $B\Gamma''$ and the proof of the theorem is concluded.

We go on now to our main theorem.

THEOREM 3. *There is a* 1-1 *corespondence between normal subgroups* G *of* Γ *such that* $\Gamma' \supset G \supset \Gamma''$, $G \neq \Gamma''$ *and ordered triplets of integers* (p, m, d) *where*

$$p > 0,\ 0 \leqq m \leqq d - 1,\ m^2 + m + 1 \equiv 0 \pmod{d}. \tag{5}$$

The correspondence arises from the coset decomposition

$$G = \sum A^k B^l \Gamma'',\ A = a^p b^{mp},\ B = b^{dp}. \tag{6}$$

Proof. Suppose that $A\Gamma''$, $B\Gamma''$ are generators of G/Γ''. It is known that if U, V are independent generators of an abelian group then so are $U^\alpha V^\beta$, $U^\gamma V^\delta$ if and only if α, β, γ, δ are integers satisfying $\alpha\delta - \beta\gamma = \pm 1$. This implies that A and B may be chosen of the form

$$A = a^p b^q, \quad B = b^r.$$

Since G is a normal subgroup of Γ, $S(A) \in G$, $S(B) \in G$. Hence integers k, l, s, t exist such that

$$S(A) \equiv A^k B^l, \quad S(B) \equiv A^s B^t.$$

If we use (2) and follow the now familiar procedure of comparing exponents, we find that $S(A) \in G$, $S(B) \in G$ if and only if integers p, m, d exist such that

$$A = a^p b^{mp}, \quad B = b^{dp}, \quad m^2 + m + 1 \equiv 0 \pmod{d}. \tag{7}$$

Since A may be replaced by A^{-1} and B by B^{-1} it is clear that both p and d may be chosen to be positive (certainly $pd \neq 0$); and since A may also be replaced by AB^k, m may be chosen modulo d.

In this fashion we obtain at least one decomposition (6) satisfying (5).

We shall show that for any given G there cannot be another decomposition (6) satisfying (5). For suppose that

$$G = \sum A_0{}^k B_0{}^l \Gamma'', \quad A_0 = a^{p_0} b^{m_0 p_0}, \quad B_0 = b^{d_0 p_0}$$

with

$$p_0 > 0,\ 0 \leqq m_0 \leqq d_0 - 1,\ m_0{}^2 + m_0 + 1 \equiv 0 \pmod{d_0}.$$

Since $A\Gamma''$, $B\Gamma''$ and $A_0\Gamma''$, $B_0\Gamma''$ are pairs of generators of Γ/Γ'' there are integers α, β, γ, δ such that $\alpha\delta - \beta\gamma = \pm 1$ and

249688

$$A_0 \equiv A^{\alpha}B^{\beta}, \quad B_0 \equiv A^{\gamma}B^{\delta}.$$

Comparing exponents we find that

$$\begin{pmatrix} p_0 & m_0p_0 \\ 0 & d_0p_0 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\begin{pmatrix} p & mp \\ 0 & dp \end{pmatrix}.$$

But $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ is unimodular and each of the matrices $\begin{pmatrix} p_0 & m_0p_0 \\ 0 & d_0p_0 \end{pmatrix}$, $\begin{pmatrix} p & mp \\ 0 & dp \end{pmatrix}$ is non-singular and in Hermite normal form. Hence they are identical and $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ is the identity matrix.

Hence we have shown that the triplets (p, m, d) and (p_0, m_0, d_0) are identical. Thus to a given group G corresponds precisely one decomposition (6) satisfying (5).

The converse is clear since conditions (5), (6) and Lemma 4 imply that G is a normal subgroup of Γ such that $\Gamma' \supset G \supset \Gamma''$ and $G \neq \Gamma''$. Furthermore G is determined by its coset decomposition modulo Γ''.

It is of interest to determine the decomposition of Γ' modulo G. We have

THEOREM 4. *G is of finite index dp^2 in Γ', and*

$$\Gamma' = \sum a^r b^s G.$$
$$0 \leqq r \leqq p-1$$
$$0 \leqq s \leqq dp-1.$$

Before proving Theorem 4, we note the interesting corollary

COROLLARY 1. *Any normal subgroup G of Γ different from Γ'' such that z is of exponent 6 modulo G is of finite index in Γ.*

Proof of Theorem 4. Since G, Γ'' are normal subgroups of Γ',

$$\Gamma'/G \cong \Gamma'/\Gamma''/G/\Gamma''. \tag{8}$$

Now Γ'/Γ'' is the free abelian group on the generators $a\Gamma''$, $b\Gamma''$ and G/Γ'' the free abelian subgroup of Γ'/Γ'' on the generators $a^p b^{mp}\Gamma''$, $b^{dp}\Gamma''$. By standard results from the theory of abelian groups and the isomorphism (8) it follows that Γ'/G is of order

$$\left| \det \begin{pmatrix} p & mp \\ 0 & dp \end{pmatrix} \right| = dp^2$$

and that $a^r b^s$, $0 \leqq r \leqq p-1$, $0 \leqq s \leqq dp-1$ form a complete set of inequivalent representatives for Γ' modulo G. The proof of the theorem is concluded.

It is of interest to note that G is *not* determined by its index in Γ' nor by the coset decomposition of Γ' modulo G. For a fixed index $(\Gamma' : G) = \mu = dp^2$, the number of possible groups G is just

$$\sum_{p^2 \mid \mu} \quad \sum_{\substack{m^2+m+1 \equiv 0 \ (\mathrm{mod}\ \mu/p^2) \\ 0 \leqq m \leqq \mu/p^2 - 1}} 1. \tag{9}$$

It is of some interest to pursue this a little further. Let $\xi(\mu)$ stand for the expression (9). Then it is not difficult to show

THEOREM 5. *Let $\xi(\mu)$ denote the number of normal subgroups of Γ of index 6μ and genus 1. Then $\xi(\mu)$ is a multiplicative arithmetic function of μ with values at the prime powers q^n given by*

$$\xi(q^n) = \begin{cases} \dfrac{1+(-1)^n}{2} & q = 2 \\ 1 & q = 3 \\ \dfrac{1+(-1)^n}{2} + \{[n/2] - (-1)^n\}(1 + (q/3)) & q > 3. \end{cases}$$

Here $(q/3)$ is the Legendre-Jacobi symbol of quadratic reciprocity.

Another interesting characterization of G is given by the corollary that follows:

COROLLARY 2. *The group G consists of all words w of Γ' such that*

$$e_a(w) \equiv 0 \ (\mathrm{mod}\, p), \quad e_b(w) \equiv m e_a(w) \ (\mathrm{mod}\, dp). \tag{10}$$

For $m = 0$, $d = 1$ (p, m, d) becomes $(p, 0, 1)$ and (10) becomes

$$e_a(w) \equiv e_b(w) \equiv 0 \ (\mathrm{mod}\, p).$$

The groups so obtained were introduced in [4] and were denoted there by $\Gamma'(p, p)$.

The principal congruence subgroup $\Gamma(6)$ is of genus 1 and of index 72 in Γ, and so of index 12 in Γ'. Thus $dp^2 = 12$ for $\Gamma(6)$. It is not difficult to show that the *only* solution of (5) for $dp^2 = 12$ is

$$(p, m, d) = (2, 1, 3).$$

Thus $\Gamma(6)$ may be characterized as the totality of words w of Γ' for which

$$e_a(w) \equiv 0 \ (\mathrm{mod}\, 2), \quad e_b(w) \equiv e_a(w) \ (\mathrm{mod}\, 6). \tag{11}$$

Restated, (11) merely says that

$$e_a(w) \equiv e_b(w) \equiv 0, 2, 4 \pmod{6}.$$

In conclusion we remark that G is a free group of rank $1 + dp^2$, since Γ' is a free group of rank 2 and $(\Gamma' \colon G) = dp^2$.

REFERENCES.

[1] R. C. Gunning, *Lectures on modular forms*, Princeton, 1962.

[2] M. Hall, Jr., *The theory of groups*, New York, 1959.

[3] A. G. Kuroš, *The theory of groups*, New York, 1955, 1956.

[4] M. Newman, "The structure of some subgroups of the modular group," *Illinois Journal of Mathematics*, vol. 6 (1962), pp. 480-487.

[5] ———, "A note on modular groups," *Proceedings of the American Mathematical Society*, vol. 14 (1963), pp. 124-125.

[6] J. Nielsen, "The commutator subgroup of the free product of cyclic groups," *Matematisk Tidsskrift. B* (1948), pp. 49-56, Danish.

ON THE INTEGRAL REPRESENTATIONS OF QUADRATIC FORMS OVER LOCAL FIELDS.

By CARL RIEHM.*

A quadratic form obtained by making a linear substitution of new variables for the variables of a given quadratic form Q, is called a *representation* of Q. One of the problems of the arithmetic theory of quadratic forms is to determine the representations of a quadratic form when the coefficient domain is an algebraic number field, the ring of integers in such a field, or a $\mathfrak{p}$-adic completion of either of them. The $\mathfrak{p}$-adic fields and the power series fields in one variable over finite fields, are usually called *local fields*.

Hasse [1] gave the solution for this problem over the rational and p-adic fields. The solution for number fields and $\mathfrak{p}$-adic fields can be found in *Introduction to Quadratic Forms* [4], by O. T. O'Meara, a book which answers the need for an account of the modern theories of equivalence and the orthogonal group, in the arithmetic case. Jones [2] gave a partial solution to the representation problem over the p-adic integers when p is odd; O'Meara [3] extended and generalized these results to include all rings of $\mathfrak{p}$-adic integers in which 2 is a unit or a prime. The present paper gives necessary and sufficient conditions for one form to be represented by a "modular" form over any ring $\mathfrak{o}$ of $\mathfrak{p}$-adic integers in which 2 is not a unit (thus 2 may be ramified); a quadratic form Q is modular if there is a field element α such that αQ has integral coefficients and unit discriminant, i. e. αQ is unimodular.

Now suppose that q and Q are two non-degenerate forms with coefficients in $\mathfrak{o}$. If q is to be represented by Q integrally, i. e. over $\mathfrak{o}$, then it must certainly be represented by Q fractionally, i. e. over the field of quotients F of $\mathfrak{o}$. Let us assume this is the case. Denote by $Q(\mathfrak{o})$ the elements represented by Q for integral values of its variables. Let Q^{*} be the form whose symmetric matrix of coefficients is inverse to that of Q. (Thus the coefficients of Q^{*}

Received November 16, 1962.
Revised January 2, 1964.

* Part of this paper was prepared while the author was a fellow at the Summer Research Institute of the Canadian Mathematical Congress.

The author is indebted to the referee for many helpful suggestions concerning the formulation of the results.

may not be integral). Suppose $\alpha^{-1}Q$ is unimodular, and let β be the greatest common divisor of the entries in the symmetric matrix of q. If q and Q have the same number of variables, the First Main Theorem (end of § 6) gives necessary and sufficient conditions for q to be represented by Q. With a few exceptions, these conditions are $\beta\mathfrak{o} \subseteq \alpha\mathfrak{o}$, $q(\mathfrak{o}) \subseteq Q(\mathfrak{o})$ and $Q^{\#}(\mathfrak{o}) \subseteq q^{\#}(\mathfrak{o})$; stronger conditions are necessary in some circumstances. From this and Theorem 6.12, we get the Second Main Theorem: Let Q be unimodular and q modular with the same (arbitrary) number of variables; then q is represented by Q if and only if $q(\mathfrak{o}) \subseteq Q(\mathfrak{o}) \subseteq q^{\#}(\mathfrak{o})$.

Now let $\nu =$ (number of variables of Q) — (number of variables of q). The Third Main Theorem (§ 7) deals with the case $\nu > 0$. If $\nu = 1$ or 2, a form q' with ν variables is determined (virtually canonically) from q and Q with the property that q is represented by Q if and only if $q + q'$ is. Then the First Main Theorem can be used to test if $q + q'$ is represented by Q. If $\nu \geqq 3$, then necessary and sufficient conditions for representation are $\beta\mathfrak{o} \subseteq \alpha\mathfrak{o}$ and $q(\mathfrak{o}) \subseteq Q(\mathfrak{o})$.

The statements of the First and Third Main Theorems do not involve the sets $q(\mathfrak{o}), Q(\mathfrak{o}), \cdots$, but rather the additive subgroups of F generated by them. These so-called norm groups are much easier to determine (both in theory and in practice), and moreover they yield "stronger" theorems in the sense that a theorem remains true if the norm groups in its statement are replaced by the sets $q(\mathfrak{o}), Q(\mathfrak{o}), \cdots$.

A result of somewhat different nature than the Main Theorems is Theorem 5.1 which describes how a representation is actually effected.

The language used in this paper is not that of quadratic forms, but rather that of the equivalent notions of quadratic spaces and lattices. §§ 1-4 introduce this language and give the preliminary results needed. §§ 5-7 contain the main results. The paper is concluded with a few remarks and partial results on the (still unsolved) problem of the representations of an *arbitrary* quadratic form Q.

Full proofs and more detailed explanations for the statements of § 1 and § 2 are to be found in [4].

1. **Basic theory.** The field F is a finite extension of the 2-adic numbers. The (unique) exponential valuation of F is written as $|\alpha|$, $(\alpha \in F)$, while the corresponding logarithmic valuation is $\operatorname{ord}\alpha$, ("order of α"). The ring of integers in F is $\mathfrak{o}$, π is a prime element (fixed from now on), $\mathfrak{p} = \pi\mathfrak{o}$ is the unique maximal ideal of $\mathfrak{o}$, and $\mathfrak{u}$ is the group of units. The residue class field $\mathfrak{o} \bmod \mathfrak{p}$ is a finite field of characteristic 2, and is therefore perfect. The

order of a fractional ideal $\mathfrak{p}^k$ is $\operatorname{ord}\mathfrak{p}^k = k$, ($k$ a rational integer). If A and B are non-zero field elements or ideals, then $A \sim B$ means that $\operatorname{ord} A \equiv \operatorname{ord} B \bmod 2$; we also use the negation of $\sim$: $A \not\sim B$, and we adopt the convention that $0 \not\sim A$ if A is non-zero. Thus for example, $1 \sim \pi^{2k}$, $\pi^{2k+1} \not\sim \mathfrak{p}^{2m}$.

If a quadratic form over F is given by means of a symmetric n by n matrix $[a_{ij}]$, then a *quadratic space* associated with this form is an n-dimensional vector space V over F provided with a bilinear form $B\colon V \times V \to F$ such that $B(x_i, x_j) = a_{ij}$ for some basis $x_1, \cdots, x_n$ of V. We set $B(x, x) = Q(x)$. To show the relationship between $[a_{ij}]$ and V, we write $V \cong \langle [a_{ij}] \rangle$ or simply $V \cong [a_{ij}]$. If S_1 and S_2 are two subsets of V, write

$$B(S_1, S_2) = \{B(x, y) \mid x \in S_1, y \in S_2\}.$$

If $B(S_1, S_2) = 0$, S_1 is *orthogonal* to S_2. An *isotropic* vector is one which is orthogonal to itself: $Q(x) = 0$. The *radical* of a space U is its subspace $\operatorname{rad} U$ consisting of those vectors x orthogonal to the whole space: $B(x, U) = 0$. A space is *regular* if its radical is zero. We shall write $V = U \perp W$ if V is the direct sum of the mutually orthogonal subspaces U and W. If $U \cong \langle \mathfrak{M} \rangle$, $W \cong \langle \mathfrak{N} \rangle$, ($\mathfrak{M}$ and $\mathfrak{N}$ symmetric matrices), we write $V \cong \langle \mathfrak{M} \rangle \perp \langle \mathfrak{N} \rangle \cong \left[\begin{array}{c|c} \mathfrak{M} & \mathfrak{O} \\ \hline \mathfrak{O} & \mathfrak{N} \end{array}\right]$. Now let S^* be the subspace V orthogonal to the subset S of V. If U is a regular subspace of the regular space V, then $V = U \perp U^*$; from this we can deduce that any space V has an orthogonal basis: $V = Fy_1 \perp \cdots \perp Fy_n$. Now define dV to be $\det[a_{ij}]$. The *discriminant* dV is uniquely determined up to multiples of non-zero squares of F. If α and β are two determinations of dV, then $\alpha = \gamma^2\beta$ for some non-zero $\gamma \in F$, and we shall write $dV \cong \alpha \cong \beta$. This notation is consistent with that introduced in § 2, i. e., $\alpha \cong \beta$ means precisely that the quadratic spaces $\langle \alpha \rangle$ and $\langle \beta \rangle$ are isometric.

A *lattice* L in a quadratic space V is a finitely generated $\mathfrak{o}$-module in V. A lattice is always free, i. e., there exist k linearly independent vectors $y_1, \cdots, y_k$ such that $L = \mathfrak{o}y_1 + \cdots + \mathfrak{o}y_k$. If we let $\mathfrak{A}$ be the k by k matrix $[B(y_i, y_j)]$, then we write $L \cong \langle \mathfrak{A} \rangle$ or simply $L \cong \mathfrak{A}$, and we define dL to be $\det \mathfrak{A}$. This time dL is uniquely determined by L up to squares of units and we adopt the conventions for it analogous to the field case. A vector $x \in L$ is *maximal* if $\pi^{-1}x \notin L$; any maximal vector x belongs to some basis of L. The subspace of V generated by L is denoted by FL; we say that L is a lattice *on* V if $FL = V$. We define $\dim L = \dim FL =$ the dimension of FL. Every line Fy of FL contains a maximal vector of L. Note that $dL \cong \alpha$ implies that $dFL \cong \alpha$. We write $L = J \perp K$ if L is a direct sum of mutually orthogonal lattices J and K; the latter are then called *components* of L, and we also write $K = J^*$; J (or K) is also said to *split* L. The *radical* of L, $\operatorname{rad} L$, is the sublattice of L consisting of all vectors orthogonal to L. We have $\operatorname{rad} L = 0$ if and only if $dL \neq 0$, and if this is the case, L is said to be *regular*. We define the *volume* of L to be $\mathfrak{v}L = dL \cdot \mathfrak{o}$. Then if K and L are regular lattices on the same space, $K \supseteq L$ implies $\mathfrak{v}K \supseteq \mathfrak{v}L$, and $K \supset L$ is equivalent to $\mathfrak{v}K \supset \mathfrak{v}L$ (if $K \supseteq L$). Finally we denote by V^α, $(\alpha \in F)$, the vector space V provided with a new bilinear form $B^\alpha(x, y) = \alpha B(x, y)$, and we say that "$V$ has been scaled by α." And L^α will be the old lattice L with the new bilinear form.

2. Isometry and representation. A linear map $\sigma \colon V \to V'$ between quadratic spaces V and V' with bilinear forms B and B' is called a *representation* if $B'(\sigma x, \sigma y) = B(x, y)$ for all $x, y \in V$. If it is one to one and onto, it is an *isometry* and we write $V \cong V'$. There are splittings $V = U \perp \operatorname{rad} V$ and $V' = U' \perp \operatorname{rad} V'$; one can then show that $V \to V'$ (i.e., there is some representation σ of V into V') if and only if $U \to U'$. Since U is regular, any representation $\sigma \colon U \to U'$ must be an isometry of U with σU, i.e., U' actually contains a replica of U. An important and basic theorem is *Witt's Theorem*: *If* $V_1 \perp V_2 \cong V_1' \perp V_2'$ *and* $V_1 \cong V_1'$, *then* $V_2 \cong V_2'$.

The problems of representation and isometry of two quadratic spaces V and V' have been completely solved over local fields. By what has been said above, it is possible to assume that V and V' are regular. The invariants which characterize the isometry class of such a space V are the obvious ones, dV and $\dim V$, plus a third, the so-called *Hasse symbol* SV. We shall not use this last invariant explicitly in this paper, and therefore will drop the matter here. However we shall use the conditions for representation, which we now list. Let $\nu = \dim V' - \dim V \geqq 0$. *Then* $V \to V'$ *if and only if*

(i) $V \cong V'$ if $\nu = 0$

(ii) $V \perp \langle dV \cdot dV' \rangle \cong V'$ if $\nu = 1$

(iii) $dV \cong -dV'$ *implies that* $F\mathfrak{H} \perp V \cong V'$ if $\nu = 2$
($V \to V'$ *always if* $\nu \geqq 3$).

Here $F\mathfrak{H}$ is the quadratic space $F\mathfrak{H} \cong \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

The definitions of representation and isometry of two lattices L and L' are the same as for quadratic spaces. Once again we can restrict our investigations to regular lattices. (See pp. 845-846 of [3]). A representation $L \to L'$ can always be extended to a representation $FL \to FL'$, so that this latter condition is a necessary one for $L \to L'$. Also if $\sigma \colon L \to L'$ is a representation, then the regularity of L implies that $L \cong \sigma L$, so that L' contains a replica of L; the converse is also true.

For any lattice L, $\mathfrak{s}L = B(L,L)$ is an ideal, the *scale* of L. Similarly the ideal $\mathfrak{n}L = Q(L) \cdot \mathfrak{o}$ is the *norm* of L. We have $2(\mathfrak{s}L) \subseteq \mathfrak{n}L \subseteq \mathfrak{s}L$. A lattice L is called *unimodular* if its associated matrix of coefficients is unimodular, (integral entries, unit determinant). L is $\mathfrak{a}$-*modular* if L^{α} is unimodular, where $\mathfrak{s}L = \mathfrak{a} = \alpha^{-1}\mathfrak{o}$. Every lattice has a *Jordan splitting* as follows:

$$L = L_1 \perp L_2 \perp \cdots \perp L_t$$

where each component L_i is modular, and $\mathfrak{s}L_1 \supset \mathfrak{s}L_2 \supset \cdots \supset \mathfrak{s}L_t$. If

$$L = K_1 \perp K_2 \perp \cdots \perp K_s$$

is any other such splitting, then $s = t$, $\dim K_i = \dim L_i$, $\mathfrak{s}L_i = \mathfrak{s}K_i$, $\mathfrak{n}L_i = \mathfrak{s}L_i \Longleftrightarrow \mathfrak{n}K_i = \mathfrak{s}K_i$, for $1 \leqq i \leqq t$.

A lattice L is called $\mathfrak{a}$-*maximal* in the space V ($\mathfrak{a}$ an ideal) if $\mathfrak{n}L \subseteq \mathfrak{a}$, and if $K \supset L$ (K another lattice in V) implies $\mathfrak{n}K \supset \mathfrak{a}$. It can be shown that $FL = V$ if L is $\mathfrak{a}$-maximal in V. Also if L is any lattice in V with $\mathfrak{n}L \subseteq \mathfrak{a}$, then there is an $\mathfrak{a}$-maximal lattice on V which contains L. The principal result concerning maximal lattices is *Eichler's Theorem*: *Any two* $\mathfrak{a}$-*maximal lattices on the same space* V *are isometric.*

The *norm group* $\mathfrak{g}L$ of a lattice L is the additive group $Q(L) + 2\mathfrak{s}L$. O'Meara introduced the norm group, and used it to characterize modular lattices: *Two* $\mathfrak{a}$-*modular lattices on the same space are isometric if and only if they have the same norm groups.* He has also determined the isometry classes of lattices in general, but we shall not need these results.

3. Norm groups and modular lattices. We shall have to know quite a bit about modular lattices, particularly about their canonical forms. This section is devoted to a summary of these facts. Most of the results will be stated for unimodular lattices—their restatements for the modular case are usually quite obvious.

A very useful concept is that of the *quadratic defect* $\mathfrak{d}(\alpha)$ of a field element α. It is defined as $\mathfrak{d}(\alpha) = \bigcap \gamma\mathfrak{o}$ where γ runs over all solutions (β, γ) of $\alpha = \beta^2 + \gamma$. It follows immediately that $\mathfrak{d}(\xi^2\alpha) = \xi^2\mathfrak{d}(\alpha)$, $\mathfrak{d}(\alpha) \subseteq \alpha\mathfrak{o}$, $\mathfrak{d}(\alpha) = \alpha\mathfrak{o}$ if and only if $\alpha \sim \pi$. One can show (63:2, [4]) that the possible defects of a *unit* are

$$0 \subset 4\mathfrak{o} \subset 4\mathfrak{p}^{-1} \subset 4\mathfrak{p}^{-3} \subset \cdots \subset \mathfrak{p}.$$

To show this, the *Local Square Theorem* (63:1, [4]) is necessary: *For any integer α, $1+4\pi\alpha$ is a square.* Also there is essentially only one unit $\Delta = 1 - 4\rho$ of defect $4\mathfrak{o}$: all other such units are of the form $\epsilon^2\Delta$, $\epsilon \in \mathfrak{u}$, (63:4, [4]). We shall fix this unit Δ, and we note in addition that ρ can be taken to be a square. Finally by 63:5 of [4], $\mathfrak{d}(\epsilon^2+\alpha) = \alpha\mathfrak{o}$ if $\epsilon \in \mathfrak{u}$, $|4| < |\alpha| < 1$, and $\alpha \sim \pi$.

An easy application of quadratic defect (63:11, [4]) is the following:

Lemma 3.1. *If V is a binary space with $dV \sim \pi$, and if α is a non-zero field element, then $Q(V)$ contains α or $\alpha\Delta$, but not both.*

Definition 3.2. A norm group $\mathfrak{g}$ is a non-zero additive subgroup of F which is bounded, i.e., $|\mathfrak{g}|$ is a bounded set of real numbers, and which admits square integers as operators: $\mathfrak{o}^2\mathfrak{g} \subseteq \mathfrak{g}$.

Throughout this paper, $\mathfrak{o}^2$ will stand for $\{\alpha^2 \mid \alpha \in \mathfrak{o}\}$, and not for the ideal product $\mathfrak{o}\cdot\mathfrak{o}$. Note that it follows directly from the definition that the sum or intersection of two norm groups is another norm group. In § 2, we defined the norm group $\mathfrak{g}L$ of a lattice L as $\mathfrak{g}L = Q(L) + 2\mathfrak{s}L$; this is easily seen to be a norm group in the sense of Definition 3.2. In fact it is the norm group generated by $Q(L)$.

If $\mathfrak{g}$ is a norm group and $\alpha \in \mathfrak{g}$, then $2\alpha\mathfrak{o} \subseteq \mathfrak{g}$. Indeed

$$2\lambda\alpha = (1+\lambda)^2\alpha - \alpha - \lambda^2\alpha \in \mathfrak{g}$$

for any $\lambda \in \mathfrak{o}$. Thus $\mathfrak{g}$ contains elements of either order parity. If $a \in \mathfrak{g}$ is an element of largest value $|a|$, (a exists since $|\mathfrak{g}|$ is a bounded discrete set), then a is called a *norm generator* of $\mathfrak{g}$. An element $b \in \mathfrak{g}$, $b \not\sim a$, with the property that $|b|$ is largest for all elements in $\mathfrak{g}$ with different order parity than a, is called a *base generator* of $\mathfrak{g}$. Since $2a\mathfrak{o} \subset \mathfrak{g}$, we have $|b| \geqq |2\pi a|$.

Lemma 3.3. *If a is a norm generator and b a base generator of the norm group $\mathfrak{g}$, then $\mathfrak{g} = a\mathfrak{o}^2 + b\mathfrak{o}$ and $b\pi^{-1}\mathfrak{o}$ is the largest ideal contained in $\mathfrak{g}$.*

Proof. Clearly $a\mathfrak{o}^2 \subset \mathfrak{g}$. Choose $\alpha \in \mathfrak{o}$ such that $|\alpha^2 a| = |b\pi^{-1}|$, and let $c \in b\pi^{-1}\mathfrak{o}$. By Lemma 3.1, we can find $\xi, \eta \in F$ such that $c \equiv \xi^2\alpha^2 a + \eta^2 b \bmod 4c\mathfrak{o}$. Since $|\xi^2\alpha^2 a|$ cannot be equal to $|\eta^2 b|$, both ξ and η must be integers, and therefore $c \in \mathfrak{g}$ since $4c\mathfrak{o} \subseteq 2b\mathfrak{o} \subset \mathfrak{g}$. Thus $a\mathfrak{o}^2 + b\mathfrak{o} \subseteq \mathfrak{g}$ and $b\pi^{-1}\mathfrak{o}$ is the largest ideal in $\mathfrak{g}$ since $b\pi^{-2} \notin \mathfrak{g}$ by choice of b. A similar argument (cf. 93:4 of [4]) shows that $\mathfrak{g} \subseteq a\mathfrak{o}^2 + b\mathfrak{o}$. And so $\mathfrak{g} = a\mathfrak{o}^2 + b\mathfrak{o}$ as required.

We now adopt the convention that whenever a norm group $\mathfrak{g}$ is expressed in the form $\mathfrak{g} = a\mathfrak{o}^2 + b\mathfrak{o}$, then a is a norm generator and b is a base generator. It should be remarked that the base generator is not quite identical to the weight generator defined by O'Meara in § 93A, [4]. For suppose that $\mathfrak{g} = \mathfrak{g}L$ is the norm group of a lattice L, and let b be a base generator. Then b is also a weight generator unless $b \in 2\pi\mathfrak{s}L$, in which case $b\pi^{-1}$ is a weight generator.

For both practical and theoretical reasons, it is important to be able to calculate the sum and intersection of two norm groups $\mathfrak{g}_1 = a_1\mathfrak{o}^2 + b_1\mathfrak{o}$, $\mathfrak{g}_2 = a_2\mathfrak{o}^2 + b_2\mathfrak{o}$. Consider first the sum. Let a be the larger in value of a_1 and a_2 and define b by $b\mathfrak{o} = \sum a^{-1}\mathfrak{d}(ac)$ where c runs over a_1, b_1, a_2, b_2. Then $\mathfrak{g}_1 + \mathfrak{g}_2 = a\mathfrak{o}^2 + b\mathfrak{o}$, (the proof is similar to that for the formula for $\mathfrak{w}$ on p. 280 of [4], and is omitted). Now the intersection. A norm generator is the largest element contained in the intersection, while a base generator is the larger in value of b_1 and b_2 (by the characterization of the largest ideal in a norm group, given in Lemma 3.3).

We now wish to discuss canonical forms. First let us write $\mathcal{A}_i(\alpha,\beta)$ for the $\mathfrak{p}^i$-modular matrix $\begin{bmatrix} \alpha & \pi^i \\ \pi^i & \beta \end{bmatrix}$; thus $\alpha, \beta \in \mathfrak{p}^i$ and $|-\pi^{2i} + \alpha\beta| = |\pi^{2i}|$. $\mathcal{A}_0(\alpha,\beta)$ is usually written $\mathcal{A}(\alpha,\beta)$. Now let L be a binary unimodular lattice, a any norm generator of L in $Q(L)$, and suppose that $dL \cong -1 + \delta$ where $\mathfrak{d}(1-\delta) = \delta\mathfrak{o}$. Then by 93:17 of [4], $L \cong \mathcal{A}(a, \delta a^{-1})$. If $K \cong \mathcal{A}_i(\alpha,\beta)$ where $\mathfrak{n}K = \alpha\mathfrak{o}$ and $\mathfrak{d}(\pi^{2i} - \alpha\beta) = \alpha\beta\mathfrak{o}$, then we shall generally write $K \cong \mathcal{B}_i(\alpha,\beta)$. Thus in the above example, $L \cong \mathcal{B}(a, \delta a^{-1})$. If ϵ is a unit, then $\mathcal{B}_i(2\pi^i\epsilon, 0) \cong \mathcal{A}_i(0,0)$ by O'Meara's theorem on modular lattices (§ 2), since their norm groups are equal (namely to $2\mathfrak{o}$), and their associated quadratic spaces are isometric. We also write $\mathcal{A}_i(0,0)$ as $\mathcal{H}_i$ and $\mathcal{H}_0$ as $\mathcal{H}$.

LEMMA 3.4. *Suppose* $L \cong \mathcal{B}(a, 2\pi\alpha)$ *where* $\alpha \in \mathfrak{o}$. *If* $c \in 2\mathfrak{o}$, *then* $L \cong \mathcal{A}(a+c, 2\pi\beta)$ *for some* $\beta \in \mathfrak{o}$.

Proof. Consider the equation $2\pi\alpha X^2 + 2X - c = 0$. If $\alpha = 0$, it has an integral solution $\xi = \frac{1}{2}c$. If $\alpha \neq 0$, it has solutions ξ, $\eta \in F$ by Hensel's Lemma (13:8, [4]) since $c \in 2\mathfrak{o}$. Moreover one of them, ξ say, must be integral since $|\xi + \eta| \geqq |\xi\eta|$. Suppose $L \cong \mathcal{A}(a, 2\pi\alpha)$ in the basis x, y. Then $L \cong \mathcal{A}(a+c, 2\pi\beta)$ in the basis $x' = x + \xi y$, $y' = y\epsilon^{-1}$, where ϵ is the unit $B(x+\xi y, y)$.

The phenomenon $|\alpha + \beta| = |\alpha|$ if $|\alpha| > |\beta|$ is often referred to as the *principle of domination*, or simply *domination*. We want to extend the use of this term to another similar situation. Namely suppose that $L \cong \mathcal{B}(a, c)$ where $ac \neq 0$, and let $\alpha = \xi^2 a + 2\xi\eta + \eta^2 c \in Q(FL)$. *Then*

$$|\alpha| = max\{|\xi^2 a|, |\eta^2 c|\}.$$

To prove this, first note that $|4| \leqq |ac| < 1$. Suppose that $|\xi^2 a| > |\eta^2 c|$; then $|4\xi^2\eta^2| \leqq |ac\xi^2\eta^2| < |\xi^2 a|^2$ whence $|2\xi\eta| < |\xi^2 a|$ and the result follows. A similar calculation proves it when $|\xi^2 a| < |\eta^2 c|$. If $|\xi^2 a| = |\eta^2 c|$, then we must have $|ac| = |4|$, i.e., $dL \cong -\Delta = -1 + 4\rho$. It is easy to check that $|\xi^2 a| = |2\xi\eta|$ also. Now suppose that

$$\xi^2 a + 2\xi\eta + \eta^2 c \equiv 0 \bmod \xi^2 a\mathfrak{p}.$$

Then there is a solution for $X^2 a + 2X + c(1 + \pi') = 0$ for some $\pi' \in \mathfrak{p}$. A straightforward calculation using the Local Square Theorem shows that the discriminant of this equation is $4\Delta\epsilon^2$ where ϵ is a unit. This is a contradiction since the discriminant is non-square, so the equation has no solution. Therefore $|\alpha| = \max\{|\xi^2 a|, |\eta^2 c|\}$ as required.

LEMMA 3.5. *If $L \cong \mathcal{B}(a, b)$ with $|b| \geqq |2|$, then*

$$L = \{x \in FL \mid Q(x) \in \mathfrak{g}L\}.$$

Proof. Let $c = \xi^2 a + 2\xi\eta + \eta^2 b \in \mathfrak{g}L$. Then $\xi \in \mathfrak{o}$ by domination. If $ab \in 4\mathfrak{o}$, then similarly $\eta \in \mathfrak{o}$. If $ab \notin 4\mathfrak{o}$, we have $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$, and so η is an integer since $c \in \mathfrak{g}L$ implies $\mathfrak{d}(ac) \subseteq ab\mathfrak{o}$.

LEMMA 3.6. *If $L \cong \mathcal{B}(a, b)$ with $ab \sim \pi$, then L contains a $b\mathfrak{p}^{-1}$-maximal lattice.*

Proof. Define $r \geqq 0$ by $|\pi^{2r+1}a| = |b|$, and suppose that $L \cong \mathcal{B}(a, b)$ in the basis x, y. Then $M = \mathfrak{o}\pi^r x + \mathfrak{o}y \subseteq L$ and $M \cong \mathcal{B}_r(\pi^{2r}a, b)$. Also $\mathfrak{g}M = \mathfrak{n}M = \pi^{2r}a\mathfrak{o}$ and $|b| \geqq |2\pi^r|$ since $|ab| > |4|$. The lemma follows by Lemma 3.5.

LEMMA 3.7. *Let V be a binary space with $dV \cong -1 + \delta \in \mathfrak{u}$, $\mathfrak{d}(1-\delta) = \delta\mathfrak{o} \supset 4\mathfrak{o}$, and let a be any field element. Then $a(1+d) \in Q(V)$ for some $d \in 4\delta^{-1}\mathfrak{o}$; and if $d \in 4\delta^{-1}\mathfrak{p}$, then $a \in Q(V)$.*

Proof. By scaling V, we can assume that $a = 1$. Of all units in $Q(V)$, let $\epsilon = 1 + d$ be that one with least quadratic defect $d\mathfrak{o}$. Suppose that $d \notin 4\delta^{-1}\mathfrak{o}$. Consider the lattice $L \cong \mathcal{B}(\epsilon, \delta\epsilon^{-1})$ on V. Since the residue class field is perfect, we can find $\xi \in F$ such that $\xi^2\delta\epsilon^{-1} \equiv -d \bmod d\mathfrak{p}$. Then the unit $\epsilon' = \epsilon + 2\xi + \xi^2\delta\epsilon^{-1}$ is in $Q(V)$ and $\mathfrak{d}(\epsilon') \subset d\mathfrak{o}$, which contradicts the choice of ϵ. This proves the first part.

To prove the last part, we assume that $d \neq 0$, $d \in 4\delta^{-1}\mathfrak{p}$. Then the unit

$$\epsilon'' = \epsilon + 2(\tfrac{1}{2}d) + (\tfrac{1}{2}d)^2\delta\epsilon^{-1}$$

is in $Q(V)$ and $\mathfrak{d}(\epsilon'') \subset d\mathfrak{o}$. This is a contradiction and so the lemma is completely proved.

3.8. *Remarks.* (i) Suppose, in the situation of Lemma 3.7, that $a \notin Q(V)$; thus $a(1+d) \in Q(V)$ for some $d \in 4\delta^{-1}\mathfrak{u}$. A close examination of the proof reveals that the reason is that

$$r = |\, ad + 2\xi + \xi^2\delta a^{-1} \,| \geqq |\, ad \,| \text{ for all } \xi \in F.$$

It follows that $r = \max\{|\, ad \,|, |\, \xi^2\delta a^{-1} \,|\}$.

(ii) A fact which will be used in conjunction with Lemma 3.7 is the following: if $\delta\mathfrak{o} = 4\mathfrak{o}$ (i.e. $dV \cong -\Delta$), then $Q(V)$ consists either of all elements of even order, or of all elements of odd order. For a proof, see 63:15, [4]. Trivially if $\delta = 0$ (i.e. $V \cong F\mathscr{H}$), then $Q(V) = F$.

(iii) Let $a \neq 0$ and $K \cong \mathscr{B}(\alpha, \beta)$ be given, with $\alpha \sim a$. By multiplying α by a suitable square unit if necessary (and β by the inverse of that unit), we can suppose that $\alpha = \pi^{2k}a + c$ where $\mathfrak{d}(\pi^{2k}a^2 + ac) = ac\mathfrak{o}$. One can show that $FK^\Delta \cong FK$, (e.g. $Q(FK^\Delta) = Q(FK)$ by 63:11a, [4]), so that it is even possible to assume that $c = 0$ if $c \in 4a\mathfrak{p}^{2k}$, since $K^\Delta \cong K$ by equality of norm groups.

LEMMA 3.9. *Let L be a ternary unimodular lattice. Then*

$$L \cong \mathscr{A}(b, 4\omega b^{-1}) \perp \langle a \rangle$$

where a is a suitable norm generator of $\mathfrak{g}L$, and b is an arbitrary base generator, and $\omega = 0$ or ρ, (ρ is as usual the unit given by $\Delta = 1 - 4\rho$).

Proof. See 93:18, [4].

Definition. A lattice is almost $\mathfrak{a}$-modular if it has a Jordan splitting $K \perp \langle \alpha \rangle$ where K is $\mathfrak{a}$-modular and $\alpha\mathfrak{o} = \pi\mathfrak{a}$. A unary (1 dimensional) almost $\mathfrak{a}$-modular lattice is a unary $\pi\mathfrak{a}$-modular lattice.

Before deriving the canonical forms for almost modular lattices, we transcribe the following useful result from 82:15 in [4].

LEMMA 3.10. *If L is a lattice containing the $\mathfrak{a}$-modular lattice J, then J splits L if and only if $B(J, L) \subseteq \mathfrak{a}$.*

LEMMA 3.11. *Let L be an almost unimodular lattice.*

(i) *If $\dim L = 2$, then $L \cong \langle a \rangle \perp \langle b \rangle$ where $a \in \mathfrak{u}$, $b \in \pi\mathfrak{u}$.*

(ii) *If* $\dim L = 3$, *then* $L \cong \mathfrak{A}(b, 4\omega b^{-1}) \perp \langle a \rangle$ *where* $a \not\sim b$, $\omega = 0$ *or* ρ, $|b| \geqq |2\pi|$, *and* $\omega = 0$ *if* $|b| = |2\pi|$.

(iii) *If* $\dim L = 4$, *then* $L \cong \mathfrak{H} \perp \langle a \rangle \perp \langle b \rangle$ *where* $a \in \mathfrak{u}$, $b \in \pi\mathfrak{u}$.

It follows that $\mathfrak{g}L = \mathfrak{o}$ *unless* $\dim L = 3$ *and* $|a| > |b|$; *in this case* $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$. *Moreover in* (ii) *if* $b_1 \in b\mathfrak{u}$, *then* $L \cong \mathfrak{A}(b_1, 4\omega b_1^{-1}) \perp \langle a \rangle$; *also* $\omega = 0$ *if and only if* FL *is isotropic.*

Proof. (i) is trivial. Consider (ii). We have $L \cong \mathfrak{B}(\alpha, \beta) \perp \langle \epsilon\pi \rangle$ where ϵ is a unit. Suppose $\alpha \sim \pi$. Choose $\xi \in \mathfrak{o}$ such that $\xi^2 \epsilon\pi \equiv -\alpha \bmod \alpha\mathfrak{p}$. Then by Lemma 3.10 and the invariant properties of the Jordan form, $L \cong \mathfrak{A}(\alpha + \xi^2\epsilon\pi, \beta) \perp \langle \epsilon'\pi \rangle$ for some unit ϵ'. We write $\mathfrak{A}(\alpha + \xi^2\epsilon\pi, \beta)$ in the form $\mathfrak{B}(\alpha', \beta')$, and if either α' or $\beta' \sim \pi$, we repeat the same process. By the results on canonical forms for binary lattices, this process eventually leads to the form given in (ii).

Now suppose that L is given in the form (ii), and define $K \cong \mathfrak{A}(b_1, 4\omega b_1^{-1}) \perp \langle a \rangle$, where $b_1 \in b\mathfrak{u}$. Clearly $FK \cong FL$. By 3.8, we can assume that $b_1 = b + \gamma$ where $\gamma \not\sim b$. If $\gamma = 0$, $K \cong L$, otherwise $\gamma \sim a$ and there is an integer η such that $\eta^2 + \gamma \equiv 0 \bmod \gamma\mathfrak{p}$. As above, the lattice

$$J \cong \mathfrak{A}(b_1 + \eta^2 a, 4\omega b_1^{-1})$$

splits K, and $dJ \cong -1 + 4\omega$ by the Local Square Theorem. Therefore $FJ \cong F\mathfrak{A}(b_1, 4\omega b_1^{-1})$ and a calculation of discriminants shows that $K \cong J \perp \langle a \rangle$. Moreover $b_1 + \eta^2 a = \delta^2(b + \gamma')$ where $\delta \in \mathfrak{u}$ and $|\gamma'| < |\gamma|$, whence $J \cong \mathfrak{A}(b_2, 4\omega b_2^{-1})$ with $b_2 = b + \gamma'$. It is now clear that after a sufficient number of applications of this technique, we can put K in the same form as that in which L was given, and so $K \cong L$. This proves the first half of the last statement of the lemma, and the other half is obvious.

Finally case (iii). We can write $L \cong \mathfrak{B}(\alpha, \beta) \perp \langle \epsilon \rangle \perp \langle \epsilon'\pi \rangle$ where ϵ and ϵ' are units. Then we can "reduce" α and β by using ϵ and $\epsilon'\pi$ as in the first part, and thus ultimately achieve the required form.

LEMMA 3.12. *Let L be unimodular of dimension $2n$, with* $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$, $\mathfrak{n}L \supset 2\mathfrak{o}$. *Then* $dL \cong (-1)^n \bmod ab\mathfrak{o}$*—that is to say, there is a determination d of dL such that* $d \equiv (-1)^n \bmod ab\mathfrak{o}$.

Proof. Since L is the orthogonal sum of binary unimodular lattices, it suffices to prove the lemma for $\dim L = 2$. But this case follows immediately from the canonical form $L \cong \mathfrak{B}(a_1, \delta a_1^{-1})$ since $\mathfrak{n}L \supset 2\mathfrak{o}$ implies that $\delta a_1^{-1} \in b\mathfrak{o}$.

LEMMA 3.13. *Let L be a quaternary unimodular lattice with* $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$. *Then*

$$L \cong \mathcal{B}(c, 4\rho c^{-1}) \perp \mathcal{B}(a, b') \text{ if } b \notin 2\mathfrak{p}$$

where c is either a suitable norm generator of $\mathfrak{g}L$, e.g., $-a+b$, or an arbitrary base generator. Also

$$L \cong \mathcal{A}(c, 0) \perp \mathcal{B}(a, b') \text{ if } b \in 2\mathfrak{p}$$

where c is any element of $\mathfrak{g}L$.

Proof. First suppose that $b \in 2\mathfrak{p}$. Then the result follows by application of 93:18, [4], and Lemma 3.4. Suppose $b \notin 2\mathfrak{p}$; we shall prove this case along with the following lemma (the converse of Lemma 3.12 when $\dim L = 4$).

LEMMA 3.14. *Let $\mathfrak{g} = a\mathfrak{o}^2 + b\mathfrak{o} \subseteq \mathfrak{o}$ be a norm group with $b \notin 2\mathfrak{p}$. Let V be a quaternary space with $\mathfrak{d}(dV) \subseteq ab\mathfrak{o}$, where dV is a unit. Then there is a quaternary unimodular lattice K on V such that $\mathfrak{g}K = \mathfrak{g}$.*

Proof. Let b' be any integer such that $1 - ab' \in \Delta \cdot dV \cdot \mathfrak{u}^2$, $\mathfrak{d}(1 - ab') = ab'\mathfrak{o}$. By field theory $F\mathcal{B}(a, b') \rightarrow V$. Write

$$V \cong \langle \alpha \rangle \perp \langle -\alpha\Delta \rangle \perp F\mathcal{B}(a, b')$$

If $\alpha \sim a$, set $c = -a + b$; if $\alpha \sim b$, set $c = b$. Then

$$K \cong \mathcal{B}(c, 4\rho c^{-1}) \perp \mathcal{B}(a, b')$$

has the required properties, since $\mathfrak{d}(dK) \subseteq ab\mathfrak{o}$ implies $b' \in b\mathfrak{o}$. And Lemma 3.13 is also proved since $K \cong L$ by equality of norm groups.

We now wish to extend the results of the last lemma to include all ternary and quaternary unimodular or almost unimodular lattices. These results are of fundamental importance for the theorems on representation which follow, and the fact that no such general conditions for the existence of binary lattices with prescribed norm group are possible, leads to considerable difficulty later on. The trouble arises because the discriminant and the norm group are too closely related in a binary lattice.

LEMMA 3.15. *The quaternary space V supports an almost unimodular lattice with norm group $\mathfrak{g}$ if and only if $\mathfrak{g} = \mathfrak{o}$, $dV \sim \pi$.*

Proof. The necessity follows by Lemma 3.11. Assume that dV is a prime element. By field theory $F\mathfrak{A} \rightarrow V$, so that $V \cong F\mathfrak{A} \perp \langle \epsilon \rangle \perp \langle -\epsilon dV \rangle$

for some unit ϵ. The lattice $L \cong \mathfrak{H} \perp \langle \epsilon \rangle \perp \langle -\epsilon dV \rangle$ fulfills the requirements of the lemma.

LEMMA 3.16. *Let V be a ternary space and $\mathfrak{g} = a\mathfrak{o}^2 + b\mathfrak{o}$ a norm group admissible as the norm group of a unimodular or almost unimodular lattice: $2\mathfrak{o} \subseteq \mathfrak{g} \subseteq \mathfrak{o}$. We further assume that dV has been expressed as a unit or a prime. Then there exists a unimodular (resp. almost unimodular) lattice L on V with $\mathfrak{g}L = \mathfrak{g}$ if and only if* (i) $dV \in \mathfrak{g}L$ (ii) $b \in 2\mathfrak{p} \Rightarrow V$ *is isotropic, and finally* (iii) $dV \sim 1$ (*resp.* $dV \sim \pi$).

Proof. The necessity is easy to prove by consideration of canonical forms. To prove the sufficiency, we first note that $-dV \in Q(V)$ is equivalent to $F\mathfrak{H} \rightarrow V$. In this case, the lattice $L \cong \mathcal{A}(c, 0) \perp \langle -dV \rangle$ satisfies the requirements, where $c = b$ unless $dV \sim \pi$ and $\mathfrak{g} = \mathfrak{o}$, in which case $c = a$. If $-dV \notin Q(V)$, then V is anisotropic and therefore $b \notin 2\mathfrak{p}$. By field theory $\langle -\Delta \cdot dV \rangle \rightarrow V$, therefore by 3.8, $V \cong \langle c \rangle \perp \langle -\Delta c \rangle \perp \langle -\Delta \cdot dV \rangle$ where c is defined above. The lattice $L \cong \mathcal{B}(c, 4\rho c^{-1}) \perp \langle -\Delta \cdot dV \rangle$ on V has the required norm group.

LEMMA 3.17. *Let L be a unimodular or almost unimodular lattice of dimension at least* 5. *Then $L \cong \mathfrak{H} \perp K$ for some lattice K.*

Proof. If L is unimodular, this is proved in 93:18 of [4]. If L is almost unimodular of even dimension, the lemma follows from Lemma 3.11, (iii); if it is of odd dimension, an argument similar to that of the latter lemma yields the result without difficulty.

LEMMA 3.18. *If $\mathfrak{H} \perp L \cong \mathfrak{H} \perp K$, then $L \cong K$.*

Proof. See 93:14 of [4].

LEMMA 3.19. *Two almost $\mathfrak{a}$-modular lattices L and K are isometric if and only if $FL \cong FK$, and $\mathfrak{g}L = \mathfrak{g}K$ if their dimension is odd.*

Proof. The necessity is obvious. We can assume that $\mathfrak{a} = \mathfrak{o}$ and that $FL = FK$. If $L \cong \mathfrak{H} \perp L'$ and $K \cong \mathfrak{H} \perp K'$, then $\mathfrak{g}L' = \mathfrak{g}K' = \mathfrak{g}L$, $FL' \cong FK'$ and it suffices to prove $L' \cong K'$. Therefore by Lemmas 3.17 and 3.11, we may assume that $\dim L = 2$ or 3. In the former case,

$$L = \{x \in FL \mid Q(x) \in \mathfrak{o}\}$$

by the principle of domination so that $K \subseteq L$ since $Q(K) \subseteq \mathfrak{o}$. Similarly $L \subseteq K$ and so $L = K$.

Suppose $\dim L = 3$. By Lemma 3.11, $L \cong \mathcal{A}(b, 4\omega b^{-1}) \perp \langle a \rangle$ and

$K \cong \mathfrak{A}(b, 4\omega' b^{-1}) \perp \langle a' \rangle$, (the same "$b$" is used in both expressions since $\mathfrak{g}L = \mathfrak{g}K$ implies their base generators are equal in value). But $\omega = \omega'$ since $\omega = 0 \Longleftrightarrow FL$ $(= FK)$ is isotropic $\Longleftrightarrow \omega' = 0$. By discriminants $\langle a \rangle \cong \langle a' \rangle$ and so $L \cong K$ as required.

4. Canonically derived lattices. In this section, L will be a lattice with the Jordan form $L = L_1 \perp \cdots \perp L_t$. We shall derive three classes of lattices from L in a canonical manner.

4.1. *The dual $L^\#$.* This lattice is defined as

$$L^\# = \{x \in FL \mid B(x, L) \subseteq \mathfrak{o}\}.$$

Most of the statements about duals which follow are easy to prove, but in any case proofs can be found in § 82F and § 82G of [4]. First $(J \perp K)^\# = J^\# \perp K^\#$ shows that $L^\# = L_1^\# \perp \cdots \perp L_t^\#$. But the dual of an $\mathfrak{a}$-modular lattice M is given by $M^\# = \mathfrak{a}^{-1}M$. Thus $L^\#$ is a lattice on FL, $K_1 \perp \cdots \perp K_t$ is a Jordan splitting of $L^\#$ where $K_i = L_{t-i+1}^\#$ and $\mathfrak{s}K_i = (\mathfrak{s}L_{t-i+1})^{-1}$. It is also easy to see that $(L^\#)^\# = L$, and that $L^\# = L$ if and only if L is unimodular. The most useful fact about duals is the following:

Principle of Duality. *If l and L are two lattices of the same dimension, then $l \to L$ if and only if $L^\# \to l^\#$.*

For suppose that $l \subseteq L$; then it follows directly from the definition that $L^\# \subseteq l^\#$, whence $L^\# \to l^\#$. Conversely, if $L^\# \to l^\#$, then $(l^\#)^\# \to (L^\#)^\#$; i. e., $l \to L$.

4.2. *The lattice $L^{\mathfrak{a}}$.* Here $\mathfrak{a}$ is an ideal, and we define

$$L^{\mathfrak{a}} = \{x \in L \mid B(x, L) \subseteq \mathfrak{a}\}.$$

(cf. § 82I, [4]). It is obviously a sublattice of L, and is equal to L if and only if $\mathfrak{s}L \subseteq \mathfrak{a}$; if $\mathfrak{a} \neq 0$, $FL^{\mathfrak{a}} = FL$. Since $(J \perp K)^{\mathfrak{a}} = J^{\mathfrak{a}} \perp K^{\mathfrak{a}}$, we have $L^{\mathfrak{a}} = L_1^{\mathfrak{a}} \perp \cdots \perp L_t^{\mathfrak{a}}$. And if M is a $\mathfrak{b}$-modular, then $M^{\mathfrak{a}} = M$ if $\mathfrak{b} \subseteq \mathfrak{a}$, $M^{\mathfrak{a}} = \mathfrak{a}\mathfrak{b}^{-1}M$ if $\mathfrak{a} \subseteq \mathfrak{b}$.

4.3. *The lattice $L_{\mathfrak{a}}$.* Here $\mathfrak{a}$ is a non-zero ideal, and we define $L_{\mathfrak{a}}$ to be the lattice whose *dual* is generated by the set of vectors

$$\{x \in L^\# \mid Q(x) \in \mathfrak{a}\}$$

Thus $L_{\mathfrak{a}}^\# \subseteq L^\#$, and so $L \subseteq L_{\mathfrak{a}}$ since clearly $FL_{\mathfrak{a}} = FL$. Also $L_{\mathfrak{a}} = L$ if $\mathfrak{n}L^\# \subseteq \mathfrak{a}$, although the converse does not hold in general; e. g., $L_{\mathfrak{a}} = L$ for any $\mathfrak{a}$ if $L \cong \mathfrak{H}$. If J is any sublattice of $L^\#$ with $\mathfrak{n}J \subseteq \mathfrak{a}$, then $J \subseteq L_{\mathfrak{a}}^\#$;

thus $L_{\mathfrak{a}}{}^{\sharp}$ is the smallest lattice containing all sublattices of $L^{\sharp}$ with norm in $\mathfrak{a}$. This also shows that *if* $\dim l = \dim L$ *and* $\mathfrak{n}L^{\sharp} \subseteq \mathfrak{a}$, *then* $l \to L$ *if and only if* $l_{\mathfrak{a}} \to L$. Indeed $L^{\sharp} \to l^{\sharp}$ if and only if $L^{\sharp} \to l_{\mathfrak{a}}{}^{\sharp}$.

In order to use this latter fact effectively, we must be able to determine $l_{\mathfrak{a}}$ from l, say from a Jordan splitting of l. Suppose that $L = J \perp K$ and $\mathfrak{n}J^{\sharp} \subseteq \mathfrak{a}$. Then by the principle of domination, it is easy to see that $L_{\mathfrak{a}} = J \perp K_{\mathfrak{a}}$. Thus the problem of determining $L_{\mathfrak{a}}$ is reduced to that of finding $K_{\mathfrak{a}}$. This is quite easy to do in the cases of importance to us, namely $\dim K = 1$ or 2. We shall now give formulas for calculating $K_{\mathfrak{a}}$ in these cases. The formulas are easily proved, usually by means of the principle of domination. To simplify the notation, we shall assume that $\mathfrak{s}K = \mathfrak{o}$, and we let $r(\alpha)$ be the smallest non-negative rational integer such that $\alpha\mathfrak{p}^{2r(\alpha)} \subseteq \mathfrak{a}$, for any non-zero $\alpha \in F$.

First suppose that K is unary: $K \cong \langle\alpha\rangle$. Then $K_{\mathfrak{a}}{}^{\sharp} \cong \langle\alpha\pi^{2r(\alpha)}\rangle$.

Next suppose that $K \cong \mathcal{B}(\alpha,\beta)$. If $\beta \neq 0$, then $K_{\mathfrak{a}}{}^{\sharp} \cong \mathcal{B}_i(\alpha\pi^{2r(\alpha)}, \beta\pi^{2r(\beta)})$, where $i = r(\alpha) + r(\beta)$. If $\beta = 0$, then the same formula holds for $K_{\mathfrak{a}}{}^{\sharp}$ unless $\alpha\pi^{2r(\alpha)} \in 2\mathfrak{p}^{r(\alpha)}$. In this case, let s be the smallest non-negative integer such that $\alpha\mathfrak{p}^{2s} \subseteq 2\mathfrak{p}^{s}$; then $K_{\mathfrak{a}}{}^{\sharp} \cong \mathcal{H}_s$.

Finally suppose that $K^{\sharp} \cong \langle\alpha\rangle \perp \langle\beta\rangle$. If $\alpha \not\sim \beta$, then

$$K_{\mathfrak{a}}{}^{\sharp} \cong \langle\alpha\pi^{2r(\alpha)}\rangle \perp \langle\beta\pi^{2r(\beta)}\rangle.$$

If $\alpha \sim \beta$ and $\beta \in \mathfrak{a} \cap \alpha\mathfrak{o}$, then the same formula holds. Suppose that $\alpha \sim \beta$ and $\mathfrak{a} \subseteq \beta\mathfrak{o} \subseteq \alpha\mathfrak{o}$. Define s by $|\alpha\pi^{2s}| = |\beta|$, and put $J^{\sharp} \cong \langle\alpha\pi^{2s}\rangle \perp \langle\beta\rangle$. Then $K_{\mathfrak{a}}{}^{\sharp} \cong J_{\mathfrak{a}}{}^{\sharp}$ and $J_{\mathfrak{a}}{}^{\sharp}$ can be calculated by the formulas of the preceding paragraph since it is modular.

The restriction of the above calculations to the unimodular case is justified by the following formulas, as are other similar assumptions in the rest of the paper. Let $\alpha \in F$ be non-zero. Then

$$(L^{\alpha})_{\mathfrak{a}}{}^{\sharp} = \alpha^{-1}(L_{\alpha\mathfrak{a}})^{\sharp}, \qquad (L^{\alpha})_{\mathfrak{a}} = L_{\alpha\mathfrak{a}}.$$

The last formula can also be written $L_{\mathfrak{a}} = (L^{\alpha})_{\alpha^{-1}\mathfrak{a}}$.

LEMMA 4.4. *If L is a binary modular lattice, then $L_{\mathfrak{a}}$ is modular.*

Proof. We can assume L is unimodular. If $L_{\mathfrak{a}}$ is not modular, then neither is $L_{\mathfrak{a}}{}^{\sharp}$, and so

$$\begin{aligned} L \supset L_{\mathfrak{a}}{}^{\sharp} &= \mathfrak{o}x \perp \mathfrak{o}y \\ &\cong \langle\alpha\rangle \perp \langle\beta\rangle \end{aligned}$$

with $|\alpha| > |\beta|$, $\alpha \sim \beta$. Since $Q(\pi^{-1}y) \in \mathfrak{a}$, y is a maximal vector of L and therefore $L = \mathfrak{o}y + \mathfrak{o}z$ where $B(y,z) = 1$. It is easily seen that $x \in \mathfrak{o}(y - \beta z)$

so that $|Q(x)| \leqq |Q(y-\beta z)| = |\beta|$, a contradiction. Thus $L_{\mathfrak{a}}$ is modular as required.

LEMMA 4.5. *If L is binary modular, then $\mathfrak{n}L_{\mathfrak{a}}^{\#} \subseteq \mathfrak{a}$ or $L_{\mathfrak{a}}^{\#} \cong \mathcal{H}_i$ for some i.*

Proof. We can assume that L is unimodular. If $\mathfrak{a} \subset \mathfrak{n}L_{\mathfrak{a}}^{\#}$, then we can find $x_1, \cdots, x_n$ in $L^{\#} = L$ such that $\mathfrak{n}L_{\mathfrak{a}}^{\#} = Q(\sum x_i)\mathfrak{o}$, (in particular $\sum x_i \in L_{\mathfrak{a}}^{\#}$), and such that $Q(x_1), \cdots, Q(x_n) \in \mathfrak{a}$. Thus it is clear that for two of the x_j's, say x_1 and x_2, we have $\mathfrak{n}L_{\mathfrak{a}}^{\#} \subseteq 2\mathfrak{p}^i$ where $B(x_1, x_2)\mathfrak{o} = \mathfrak{p}^i$. Since $Q(x_1)$ and $Q(x_2)$ are in $\mathfrak{a} \subset 2\mathfrak{p}^i$, the Local Square Theorem shows that $dK \cong -\pi^{2i}$ where $K = \mathfrak{o}x_1 + \mathfrak{o}x_2$. Since $\mathfrak{n}K = 2\mathfrak{p}^i$, we actually have $K \cong \mathcal{H}_i$. Now $\mathfrak{s}K \supseteq \frac{1}{2}\mathfrak{n}L_{\mathfrak{a}}^{\#} \supseteq \mathfrak{s}L_{\mathfrak{a}}^{\#}$ so that $L_{\mathfrak{a}}^{\#} = K$ since it is obvious that $K \subseteq L_{\mathfrak{a}}^{\#}$. Therefore $L_{\mathfrak{a}}^{\#} \cong \mathcal{H}_i$ as required.

5. Existence of certain sublattices.

The first result, Theorem 5.1, is important for two reasons. First, it is used to prove Lemma 5.2, which in turn is used constantly throughout the rest of the paper to construct lattices with given norm groups. Secondly, it shows the manner in which any sublattice l of a given lattice L is imbedded in L. For example, suppose that L is unimodular, and that $l = l_1 \perp \cdots \perp l_t$ is a Jordan splitting. Then there is an (almost) modular lattice $L_1 \subseteq L$ with scale $\mathfrak{s}l_1$ such that $L_1 = l_1 \perp L_1'$; again there is an (almost) modular lattice $L_2 \subseteq L_1'$ with scale $\mathfrak{s}l_2$ such that $L_2 = l_2 \perp L_2'$. This process can be continued until the lattice l is ultimately constructed.

THEOREM 5.1. *Suppose that the lattice l is represented by another lattice L. Let $\mathfrak{a}$ be an ideal contained in $(\mathfrak{s}L^{\#})^{-1}$. Suppose $l = l_1 \perp l_2$ where $\mathfrak{s}l_1^{\#} \subseteq \mathfrak{a}^{-1}$, $\mathfrak{s}l_2 \subseteq \mathfrak{a}$ and $\dim l_1 < \dim L$. Then there exists an $\mathfrak{a}$-modular or almost $\mathfrak{a}$-modular lattice k such that $l_2 \rightarrow k$, $l_1 \perp k \rightarrow L$, $\dim(l_1 \perp k) = \dim L$.*

Remarks. It is possible that l_1 or $l_2 = 0$. The condition $\mathfrak{a} \subseteq (\mathfrak{s}L^{\#})^{-1}$ means that $\mathfrak{a}$ is contained in the scales of all Jordan components of L, while $\mathfrak{s}l_1^{\#} \subseteq \mathfrak{a}^{-1}$ has a similar meaning for l_1 if $l_1 \neq 0$.

Proof. We can assume that $l \subseteq L$. If $\dim l < \dim L$, say $FL = Fl \perp V$, choose an orthogonal basis for V and let l' be the lattice spanned by them. For suitably large m, $\mathfrak{p}^m l' \subseteq L$ and $\mathfrak{s}(\mathfrak{p}^m l') \subseteq \mathfrak{a}$. It is clear that adjoining $\mathfrak{p}^m l'$ to l allows us to assume that $\dim l = \dim L$. In particular $l_2 \neq 0$.

Suppose that l_2 is not (almost) $\mathfrak{a}$-modular. Then there is a maximal vector $x \in l_2$ such that $B(x, l_2) \subset \mathfrak{a}$ and $Q(x) \in \mathfrak{a}\mathfrak{p}^2$, (consider a Jordan splitting of l_2). If x is not a maximal vector *in* L, then $\pi^{-1}x \in L$ and we set $J = \mathfrak{o}(\pi^{-1}x) + l_2$. Then

$$\mathfrak{s}J = B(J, J) = \pi^{-2}Q(x) \cdot \mathfrak{o} + B(\pi^{-1}x, l_2) + \mathfrak{s}l_2 \subseteq \mathfrak{a}.$$

Now suppose that x *is* a maximal vector in L. By considering any orthogonal splitting of L into unary and binary modular lattices, we see that there is a vector y in L such that

$$B(x, y)\mathfrak{o} = B(y, L) = \mathfrak{a}.$$

Let $y = y_1 + y_2$ where $y_1 \in Fl_1$ and $y_2 \in Fl_2$. Then

$$B(y_1, l_1) = B(y, l_1) \subseteq B(y, L) = \mathfrak{a},$$

whence $y_1 \in l_1 \subset L$, and therefore $y_2 \in L$. Also $B(x, y_2)\mathfrak{o} = B(x, y)\mathfrak{o} = \mathfrak{a}$, so that $y_2 \notin l_2$ (since $B(x, l_2) \subset \mathfrak{a}$). Moreover if $J = \mathfrak{o}y_2 + l_2$, then $\mathfrak{s}J \subseteq \mathfrak{a}$ by an argument like the one above—note that $Q(y_2) = Q(y) - Q(y_1)$ and $Q(y) \in \mathfrak{a}$.

Using the original vector x, we have found a lattice $J \supset l_2$ with $\mathfrak{s}J \subseteq \mathfrak{a}$. Thus letting $\dim l_2 = m$, we have $\mathfrak{v}l_2 \subset \mathfrak{v}J \subseteq \mathfrak{a}^m$. If J is not (almost) $\mathfrak{a}$-modular, we repeat the process. Continuing in this manner, we get a chain of lattices on Fl_2 with volume at most $\mathfrak{a}^m$:

$$l_2 \subset J \subset J' \subset \cdots \subseteq L,$$

$$\mathfrak{v}l_2 \subset \mathfrak{v}J \subset \mathfrak{v}J' \subset \cdots \subseteq \mathfrak{a}^m.$$

Since the volumes are bounded by $\mathfrak{a}^m$, the chain must break off after a finite number of steps, i.e., we eventually get the desired lattice k.

LEMMA 5.2. *Suppose l and L are lattices such that* $\dim l < \dim L$, *$Fl \to FL$ and $\mathfrak{g}l \subseteq \mathfrak{g}L$. Let $\mathfrak{a}$ be an ideal such that $\mathfrak{s}l^\#$ and $\mathfrak{s}L^\#$ are contained in $\mathfrak{a}^{-1}$. Then there is an $\mathfrak{a}$-modular or almost $\mathfrak{a}$-modular lattice k such that $F(l \perp k) \cong FL$ and $\mathfrak{g}k \subseteq \mathfrak{g}L$.*

Proof. Let $\mathfrak{s}l = \mathfrak{p}^s$. If $l' \cong \langle\alpha\rangle$ is a unary component of l, it is clear that $l' \to \mathcal{A}_s(\alpha, 0) \cong L'$; if $l'' \cong \mathcal{A}_j(\alpha, \beta)$ is a binary component of l, then $l'' \to \mathcal{A}_s(\alpha, 0) \perp \mathcal{A}_s(\beta, 0) \cong L''$. Clearly $\mathfrak{g}L'$ and $\mathfrak{g}L''$ are both in $\mathfrak{g}L$. Thus expressing l as an orthogonal sum of unary and binary modular lattices, we can find a $\mathfrak{p}^s$-modular lattice M such that $l \to L \perp M$, $\mathfrak{g}M \subseteq \mathfrak{g}L$, and such that FM is a hyperbolic space, i.e., an orthogonal sum of spaces isometric to $F\mathcal{H}$. Now apply Theorem 5.1 to find an (almost) $\mathfrak{a}$-modular lattice K such that $\mathfrak{g}K \subseteq \mathfrak{g}L$ and $F(l \perp K) \cong F(L \perp M)$. By Lemma 3.17, $K = K' \perp M'$ where $\dim K' \leqq 4$ and FM' is zero or a hyperbolic space.

By removing an equal number of hyperbolic planes $F\mathcal{H}$ from FM and FM' and applying Witt's Theorem, we can assume that either M or M' is zero. If $M=0$, then we can choose $k=K$. Suppose $M\neq 0$, $M'=0$, so that $\dim K\leqq 4$. Since $\dim l<\dim L$, we must have $FM\cong F\mathcal{H}$. Now $FL\cong Fl\perp V$ for some space V, whence

$$Fl\perp V\perp F\mathcal{H}\cong Fl\perp FK;$$

therefore FK is isotropic. Choose a maximal isotropic vector x in K and another vector y in K such that $B(x,y)\mathfrak{o}=\mathfrak{a}$; this is clearly possible if K is modular; if K is almost modular their existence follows immediately from the canonical forms of Lemma 3.11. Then by Lemma 3.10, $K=(\mathfrak{o}x+\mathfrak{o}y)\perp k$, and k is the required lattice.

LEMMA 5.3. *If L is a maximal lattice, then $l\to L$ if $\mathfrak{n}l\subseteq\mathfrak{n}L$ and $Fl\to FL$.*

Proof. Assume $Fl\subseteq FL$. Let M be a $\mathfrak{n}L$-maximal lattice on FL containing l (cf. § 2). Then $l\to M\cong L$ by Eichler's Theorem, as required.

The last lemma can sometimes be used in conjunction with the next to test for representation.

LEMMA 5.4. *A lattice L contains an $\mathfrak{a}$-maximal lattice (on FL) if $\mathfrak{a}\subseteq 4(\mathfrak{n}L^{\#})^{-1}$.*

Proof. We shall prove the lemma for $\mathfrak{a}=4(\mathfrak{n}L^{\#})^{-1}$; the more general case follows from this by Lemma 5.3 and the fact that a given space supports a $\mathfrak{b}$-maximal lattice for every $\mathfrak{b}$. Let M be a $(\mathfrak{n}L^{\#})$-maximal lattice containing $L^{\#}$. Then $\mathfrak{s}M\subseteq\frac{1}{2}\mathfrak{n}M=\frac{1}{2}\mathfrak{n}L^{\#}$ and so

$$M'=2(\mathfrak{n}L^{\#})^{-1}M\subseteq(\mathfrak{s}M)^{-1}M\subseteq M^{\#}\subseteq L.$$

M' is clearly the required lattice.

COROLLARY 5.5. *A modular lattice L with $\mathfrak{n}L=2\mathfrak{s}L$ is $\mathfrak{n}L$-maximal. In particular $\mathcal{H}_i$ and $\mathcal{A}_i(2\pi^i,2\rho\pi^i)$ are $2\mathfrak{p}^i$-maximal.*

6. Representation when the dimensions are equal.

In this section we shall give conditions under which l is represented by a unimodular lattice L of the same dimension.

LEMMA 6.1. *Let L be a ternary unimodular (resp. almost unimodular) lattice and l be a ternary almost $\mathfrak{p}$-modular (resp. $\mathfrak{p}$-modular) lattice. Then $l\to L$* if *$Fl\cong FL$ and $\pi^2\mathfrak{g}L\subseteq\mathfrak{g}l\subseteq\mathfrak{g}L$.*

Proof. Put the two lattices in the canonical forms of Lemma 3.9 and 3.11:

$$L \cong \mathcal{A}(b, 4\omega b^{-1}) \perp \langle a\rangle, \qquad l \cong \mathcal{A}_1(\beta, 4\omega' \beta^{-1}\pi^2) \perp \langle \alpha\rangle.$$

Now $\omega = \omega'$ since otherwise one of FL and Fl would be isotropic and the other not. Thus by discriminants and the fact that $\mathfrak{g}l \subseteq \mathfrak{g}L$, we have $\alpha \in a\mathfrak{o}^2$. Since $\alpha \sim a$, we must have $\beta \sim b$ and therefore $\beta \in b\mathfrak{o}$. Since β can be replaced by any element of the same value (by 3.9 or 3.11), we can assume $\beta = \pi^{2k}b$. But since $\pi^2\mathfrak{g}L \subseteq \mathfrak{g}l$, $k = 0$ or 1 and the reader can easily check that $\mathcal{A}_1(\beta, 4\omega\beta^{-1}\pi^2) \rightarrow \mathcal{A}(b, 4\omega b^{-1})$ in either of these cases. Since $\langle \alpha\rangle \rightarrow \langle a\rangle$ by what was said above, $l \rightarrow L$ as required.

LEMMA 6.2. *Let l and L be almost modular lattices of the same even dimension. Then $l \rightarrow L$ if $\mathfrak{s}l \subseteq \mathfrak{s}L$ and $Fl \cong FL$.*

Proof. By Lemmas 3.17 and 3.11, $L = L' \perp K$, $l = l' \perp k$ where L' and l' are zero or modular hyperbolic lattices of the same dimension (a hyperbolic lattice is the orthogonal sum of $\mathcal{H}_i$'s), and K and k are binary almost modular lattices. Clearly $l' \rightarrow L'$. By Witt's Theorem $Fk \cong FK$. By the principle of domination K is maximal. Therefore $k \rightarrow K$ by Lemma 5.3 since $\mathfrak{n}k = \mathfrak{s}l \subseteq \mathfrak{s}L = \mathfrak{n}K$.

LEMMA 6.3. *Let $\mathfrak{g}$ be a norm group contained in another norm group $\mathfrak{g}'$. Then $\mathfrak{g}$ is a proper subgroup of $\mathfrak{g}'$ if and only if either some norm generator of $\mathfrak{g}$ or every base generator of $\mathfrak{g}$ is in $\pi^2\mathfrak{g}'$.*

The proof is straightforward and is omitted.

LEMMA 6.4. *Suppose L is unimodular, l is $\mathfrak{p}$-modular and* $\dim L = \dim l = 4$. *Then $l \rightarrow L$ if $Fl \cong FL$ and $\pi^2\mathfrak{g}L \subseteq \mathfrak{g}l \subseteq \mathfrak{g}L$.*

Proof. If $\mathfrak{n}L = 2\mathfrak{o}$, then $l \rightarrow L$ by Corollary 5.5 and Lemma 5.3. We can therefore assume that $\mathfrak{n}L \supset 2\mathfrak{o}$. We put $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$. Note that no norm or base generator of $\mathfrak{g}l$ can be in $\pi^4\mathfrak{g}L$ since otherwise it would be in $\pi^2\mathfrak{g}l$ which is impossible by Lemma 6.3. It is clear that there are three cases to be considered: $\mathfrak{n}l$ equal to $\mathfrak{n}L$, $\pi\mathfrak{n}L$ or $\pi^2\mathfrak{n}L$.

First Case: $\mathfrak{n}l = \mathfrak{n}L$. A suitable choice of a is a norm generator of $\mathfrak{g}l$ and either b or $\pi^2 b$ is a base generator for $\mathfrak{g}l$. By Lemma 3.13,

$$L \cong \mathcal{B}(c, 4\omega c^{-1}) \perp \mathcal{B}(a, b')$$

where $c = b$ or $-a + b$, $\omega = 0$ if $b \in 2\mathfrak{p}$, otherwise $\omega = \rho$. If b is a base generator of $\mathfrak{g}l$, (i.e., $\mathfrak{g}l = \mathfrak{g}L$), it is clear that

$$L \leftarrow \mathcal{B}_1(c, 4\omega c^{-1}\pi^2) \perp \mathcal{B}_1(a, b'\pi^2) \cong l;$$

the isometry results from the equality of norm groups, and O'Meara's theorem on the isometry of modular lattices. Similarly if $\pi^2 b$ is a base generator of $\mathfrak{g}l$, then

$$L \leftarrow \mathcal{A}_1(\pi^2 c, 4\omega c^{-1}) \perp \mathcal{B}_1(a, b'\pi^2) \cong l,$$

unless $\omega = \rho$ and $|b| = |c| = |2|$ since in this situation, $4\omega c^{-1} \notin \mathfrak{g}l$. However we can then apply Lemma 3.13 to l:

$$\begin{aligned} l &\cong \mathcal{A}_1(2\pi^2, 0) \perp \mathcal{B}_1(a, \beta) \\ &\rightarrow \mathcal{B}(2, 0) \perp \mathcal{B}(a, \beta\pi^2) \cong L \end{aligned}$$

by equality of norm groups. Note that $\beta\pi^{-2} \in \mathfrak{g}L$ since $\beta \in 2\mathfrak{p}^2$.

Second Case: $\mathfrak{n}l = \pi\mathfrak{n}L$. It is easily seen that $\mathfrak{g}l = b\mathfrak{o}^2 + \pi^2 a\mathfrak{o} = \mathfrak{n}l$, $\mathfrak{g}L = \mathfrak{n}L$. Let x_1, x_2, x_3, x_4, be a basis for L corresponding to the canonical form in the previous case. Suppose that $c = -a + b$. Then the lattice

$$J = \mathfrak{o}(x_1 + x_3) + \mathfrak{o}x_2 \cong \mathcal{A}(b, 4\omega c^{-1})$$

splits L by Lemma 3.10: $L = J \perp K$; moreover $\mathfrak{n}K = \mathfrak{n}L$ since $\mathfrak{n}J \subset \mathfrak{n}L$. Therefore for some norm generator a' of $\mathfrak{g}L$ and for either possibility for c (i.e., $-a + b$ or b), we have

$$\begin{aligned} L &\cong \mathcal{A}(b, 4\omega c^{-1}) \perp \mathcal{B}(a', b'') \\ &\leftarrow \mathcal{A}_1(b, 4\omega c^{-1}\pi^2) \perp \mathcal{A}_1(a'\pi^2, b'') \cong l; \end{aligned}$$

the assumption that $nL \supset 2\mathfrak{o}$ ensures that $b'' \in a'\mathfrak{p} = \mathfrak{g}l$.

Third Case: $\mathfrak{n}l = \pi^2\mathfrak{n}L$. Define $k = (l^\#)^\pi$ and $K = (L^\#)^\pi$; k is unimodular and K is $\mathfrak{p}$-modular. An easy computation shows that $\mathfrak{g}K = \pi\mathfrak{g}L$ and $\mathfrak{g}k = \pi^{-1}\mathfrak{g}l$. Therefore $\mathfrak{n}K = \mathfrak{n}k$ and the given norm group condition is equivalent to $\pi^2\mathfrak{g}k \subseteq \mathfrak{g}K \subseteq \mathfrak{g}k$. Applying the results of the First Case, we get $K \rightarrow k$, i.e., $(L^\#)^\pi \rightarrow (l^\#)^\pi$ whence $L^\# \rightarrow l^\#$ and finally $l \rightarrow L$ as required.

Definition 6.5. A splitting $l = l_1 \perp \cdots \perp l_t$ is called a quasi-Jordan splitting of l if $l_1, \cdots, l_{t-1}$ are modular, and l_t is modular or almost modular, $\mathfrak{s}l_1 \supset \mathfrak{s}l_2 \supset \cdots \supset \mathfrak{s}l_t$, and with the additional proviso that if $\dim l_t = 1$, then l_t is modular and moreover $\mathfrak{s}l_t \subset \pi\mathfrak{s}l_{t-1}$ if $t > 1$.

Any lattice l has a quasi-Jordan splitting—it can be derived directly from a Jordan splitting. And it is clear that t, $\dim l_i$ and $\mathfrak{s}l_i$, $(1 \leqq i \leqq t)$, are the same for every quasi-Jordan splitting of l.

The conditions $\mathfrak{s}l \subseteq \mathfrak{s}L$, $\mathfrak{g}l \subseteq \mathfrak{g}L$, and $Fl \rightarrow FL$ are obvious necessary conditions for $l \to L$. If $\dim l = \dim L$, then $L^{\#} \to l^{\#}$ yields the additional necessary condition $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}l^{\#}$; if L is unimodular or almost unimodular, then $L \subseteq L^{\#}$ and so $\mathfrak{g}L \subseteq \mathfrak{g}l^{\#}$ is also necessary. This condition is usually not quite strong enough, and it is often strengthened, as in the next theorem, by using one of the derived lattices $l_{\mathfrak{a}}$ introduced in 4.3.

THEOREM 6.6. *Let L be unimodular or almost unimodular, and let l be another lattice of the same dimension. Put $\mathfrak{s} = (\mathfrak{s}l_t)^{-1}$ and suppose that $\dim l_t \geqq 3$, where l_t is the last component in a quasi-Jordan splitting of l. Then $l \to L$ if and only if $Fl \cong FL$, $\mathfrak{s}l \subseteq \mathfrak{s}L$, and $\mathfrak{g}l \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{s}}^{\#}$.*

Remark. If l_t is $\mathfrak{p}^i$-modular, then $l_{\mathfrak{s}}^{\#} = l_t^{*\#} \perp (\pi^{-i}l_t) = l^{\#}$. If l_t is almost modular, the principle of domination shows that $l_{\mathfrak{s}}^{\#} = l_t^{*\#} \perp (\pi^{-i}l_t) \subset l^{\#}$.

Proof. If $l \to L$, then $L \to l^{\#}$. Since $\mathfrak{n}L \subseteq \mathfrak{s}$, it follows from 4.3 that $L \to l_{\mathfrak{s}}^{\#}$. The necessity is now obvious.

Now let us consider the sufficiency. The proof is by induction on $\tau(l) = \operatorname{ord}(\mathfrak{s}l_t) \geqq 0$. If $\tau(l) = 0$, then l and L are both unimodular or almost unimodular with equal norm groups (since $l_{\mathfrak{s}}^{\#} = l$), and so $l \cong L$. Hence we may assume that $i = \tau(l) > 0$. We may also assume that $\mathfrak{n}L \supset 2\mathfrak{o}$ since otherwise $l \to L$ by 5.5 and 5.3.

Now suppose that $\dim l_t > 4$. By Lemma 3.17, $l_t \cong l_t' \perp \mathcal{H}_i$, and $L \cong L' \perp \mathcal{H}$. Clearly l_t' is the last component in a quasi-Jordan splitting of $l' \cong l_t^* \perp l_t'$, and $\mathfrak{s}l_t' = \mathfrak{s}l_t$. Also $\mathfrak{g}l = \mathfrak{g}l'$, $\mathfrak{g}L' = \mathfrak{g}L$, $\mathfrak{g}l_{\mathfrak{s}}'^{\#} = \mathfrak{g}l_{\mathfrak{s}}^{\#}$, and $Fl' \cong FL'$ by Witt's Theorem. Since it is clear that $l \cong l' \perp \mathcal{H}_i \to L \cong L' \perp \mathcal{H}$ if $l' \to L'$, it suffices to show that $l' \to L'$. This allows us to assume that $\dim l_t = 3$ or 4.

Put $\mathfrak{t} = \mathfrak{s}l_t$ and $\mathfrak{g}l^{\mathfrak{t}} = \alpha\mathfrak{o}^2 + \beta\mathfrak{o}$. By 93:13 of [4],

$$l \perp \mathcal{H}_i \perp \mathcal{H}_i \cong l \perp \mathcal{A}_i(\alpha, 0) \perp \mathcal{A}_i(\beta, 0).$$

By Lemma 3.17,

$$l_t \perp \mathcal{A}_i(\alpha, 0) \perp \mathcal{A}_i(\beta, 0) \cong l_t' \perp \mathcal{H}_i \perp \mathcal{H}_i$$

for some (almost) $\mathfrak{t}$-modular lattice l_t'; it is clear that $\mathfrak{g}l_t' = \mathfrak{g}l^{\mathfrak{t}}$ and $Fl_t' \cong Fl_t$. Using Lemma 3.18, we get $l \cong l_t^* \perp l_t'$—that is to say, there is a quasi-Jordan splitting in which the norm group of the last component is equal to $\mathfrak{g}l^{\mathfrak{t}}$. Let us assume that the splitting $l = l_1 \perp \cdots \perp l_t$ already has this property. Since $l_{\mathfrak{s}}^{\#} = \pi^{-i}l^{\mathfrak{t}}$, this means that $\mathfrak{g}l_{\mathfrak{s}}^{\#} = \pi^{-2i}\mathfrak{g}l_t$.

Now define a norm group $\mathfrak{g} = \mathfrak{g}L \cap \mathfrak{p}^{i-1} \cap \pi^{-2}\mathfrak{g}l_t$. We shall find an (almost) $\mathfrak{p}^{i-1}$-modular lattice k on Fl_t with $\mathfrak{g}k = \mathfrak{g}$. Suppose for the moment that this has been done. Then we have $l_t \to k$ by 6.1, 6.2 or 6.4—the condition in Lemma 6.1 or 6.4 involving norm groups becomes $\mathfrak{g}l_t \subseteq \mathfrak{g}k \subseteq \pi^{-2}\mathfrak{g}l_t$ in this situation—and $\tau(J) < \tau(l)$ where $J = l_t^* \perp k$. Clearly $\mathfrak{g}J \subseteq \mathfrak{g}L$. Also $\mathfrak{g}L \subseteq \mathfrak{g}J_{\mathfrak{s}'}^{\#}$ where $\mathfrak{s}' = (\mathfrak{s}k)^{-1}$ since

$$\begin{aligned} \mathfrak{g}J_{\mathfrak{s}'}^{\#} &\supseteq \mathfrak{g}(\pi^{-(i-1)}k) = \pi^{-2(i-1)}\mathfrak{g}k = (\pi^{-2(i-1)}\mathfrak{g}L) \cap \mathfrak{p}^{-i+1} \cap \pi^{-2i}\mathfrak{g}l_t \\ &\supseteq \mathfrak{g}L \cap \pi^{-2i}\mathfrak{g}l_t = \mathfrak{g}L \cap \mathfrak{g}l_{\mathfrak{s}}^{\#} = \mathfrak{g}L \end{aligned}$$

Therefore $J \to L$ by induction, whence $l \to J \to L$. The theorem will therefore be proved as soon as we show the existence of k.

By Lemma 5.2, there is an (almost) $\mathfrak{p}^{i-1}$-modular lattice K on Fl_t with $\mathfrak{g}K \subseteq \mathfrak{g}L$, unless L is almost unimodular and $i = 1$. But such a lattice K also exists in this last case. To show this, first notice that we can find a unimodular lattice M on a hyperbolic space of dimension $2n = 2 \dim l_t^*$ such that $l_t^* \to M$ and $\mathfrak{g}M \subseteq \mathfrak{g}L$ by the same technique used at the beginning of the proof of Lemma 5.2 (we may suppose that $l_t^* \neq 0$ since otherwise we can take $K = L$). Thus $l_t^* \to L \perp M$ and so $L \perp M \cong l_t^* \perp K'$ by Lemma 3.10. Now $\dim K' \geqq 3 + \dim M$, whence K' is split by a unimodular lattice M' consisting of an orthogonal sum of hyperbolic planes with $\dim M' = \dim M$, by Lemma 3.17. Define K by $K' = K \perp M'$. Then $\mathfrak{g}K \subseteq \mathfrak{g}K' \subseteq \mathfrak{g}(L \perp M) = \mathfrak{g}L$ and $FK \cong Fl_t$ by Witt's Theorem. Thus K is the required lattice.

First Case: $\dim l_t = 4$ *and* l_t *is modular.* First suppose that l_t is $\mathfrak{p}$-modular ($i = 1$, $t = 1$ or 2), and that $b \in 2\mathfrak{p}$ where $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$. Then $K \cong \mathcal{H} \perp K'$ where K' is binary unimodular, by the second part of Lemma 3.13. If $\mathfrak{n}l_t = 2\mathfrak{p} = \mathfrak{g}l_t$, then $l_t \cong \mathcal{H}_1 \perp \mathcal{H}_1$ or $\mathcal{H}_1 \perp \mathcal{B}_1(2\pi, 2\rho\pi)$ since any base generator of $\mathfrak{g}l_t$ is in $2\mathfrak{p}^2$. In the first possibility $k \cong \mathcal{H} \perp \mathcal{H}$ has the required properties. The second possibility is ruled out by the existence of K; namely $FK' \cong F\mathcal{B}_1(2\pi, 2\rho\pi)$, and from this and 3.8(ii), it is clear that FK' cannot support a unimodular lattice whose norm group is in $\mathfrak{g}L$. Now suppose that $\mathfrak{n}l_t \supset 2\mathfrak{p}$. Since K' is unimodular and $\mathfrak{g}K' \subseteq \mathfrak{g}L$, it is clear that 2π is a base generator for $\mathfrak{g}K'$ as well as for $\mathfrak{g}L$. Applying 3.8 (iii) and 3.4 to K', we get $K' \cong \mathcal{B}(\pi^{2m}a, \beta)$—if $\mathfrak{n}K' = 2\mathfrak{o}$, this follows directly from 3.8(ii) and equality of norm groups. There is a lattice on FK' with matrix $\mathcal{B}(a, \beta\pi^{2m})$, and its norm group is clearly in $\mathfrak{g}L$; we may therefore suppose that $m = 0$, i.e. $K' \cong \mathcal{B}(a, \beta)$ for suitable β. Now $\mathfrak{g}l_t \subseteq \mathfrak{g}L$ and $\mathfrak{n}l_t \supset 2\mathfrak{p}$ implies that $\mathfrak{n}l_t \sim a$, whence 2π is a base generator of $\mathfrak{g}l_t$, whence there is a norm generator of $\mathfrak{g}l_t$ of the form $\pi^{2n}a$ for some $n \geqq 0$. We say that

$$l_t \cong \mathcal{H}_1 \perp \mathcal{B}_1(\pi^{2n}a, \beta\pi^{-2n+2}).$$

Since $Fl_t \cong FK$, it is clear that we need only show that $\beta\pi^{-2n+2} \in \mathfrak{g}l_t$ since then the two norm groups will be equal. To this end we notice that $Fl_t \cong FK$

is isotropic, and we may therefore split a binary isotropic lattice h off l_t; if h^* is the orthogonal complement of h in l_t, then h^* is a binary $\mathfrak{p}$-modular lattice on FK' with $\mathfrak{g}h^* \subseteq \mathfrak{g}l_t$. A comparison of a canonical form of h^* and $\mathcal{B}_1(\pi^{2n}a, \beta\pi^{-2n+2})$ shows that $\beta\pi^{-2n+2} \in \mathfrak{g}l_t$ as required. Now an easy calculation shows that

$$\mathfrak{g} = \begin{cases} a\mathfrak{o}^2 + 2\pi\mathfrak{o} & \text{if } n = 0, \\ \pi^{2n-2}a\mathfrak{o}^2 + 2\pi\mathfrak{o} & \text{if } n > 0, \end{cases}$$

and therefore that a proper choice of k is

$$k \cong \begin{cases} \mathcal{H} \perp \mathcal{B}(a, \beta) & \text{if } n = 0, \\ \mathcal{H} \perp \mathcal{B}(\pi^{2n-2}a, \beta\pi^{-2n+2}) & \text{if } n > 0; \end{cases}$$

the reader can easily check that β or $\beta\pi^{-2n+2}$ as the case may be, is in $\mathfrak{g}$.

From now on (in the First Case), we may suppose that either $i \geqq 2$, or that $i = 1$ and $|b| \geqq |2|$; this implies that $\beta_0 \notin 2\mathfrak{p}^i$, where $\mathfrak{g} = \alpha_0\mathfrak{o}^2 + \beta_0\mathfrak{o}$, (directly from the definition of $\mathfrak{g}$). Thus writing $dFl_t \cong 1 - \delta$ where $\mathfrak{d}(1 - \delta) = \delta\mathfrak{o}$, we see by Lemma 3.14 that k exists if

$$\delta \in \alpha_0\beta_0\mathfrak{p}^{-2(i-1)}. \tag{1}$$

Let us write $\mathfrak{g}l_t = \alpha\mathfrak{o}^2 + \beta\mathfrak{o}$. If $\mathfrak{g}l_t \subset \mathfrak{g}$, then Lemma 6.3 shows that (1) is satisfied since

$$\delta \in \alpha\beta\mathfrak{p}^{-2i} + 4\mathfrak{o}$$

in any case. Assume that $\mathfrak{g} = \mathfrak{g}l_t$. Then we must have $|\beta| = |b|$ and $|\alpha| = |a|$ or $|\pi^i|$. Therefore $\mathfrak{g} = \mathfrak{g}L \cap \mathfrak{p}^{i-1}$, and so $\mathfrak{g}K \subseteq \mathfrak{g}$, and (1) results from the similar expression for $\mathfrak{g}K$.

Second Case: $\dim l_t = 4$ *and* l_t *is almost modular.* It is easily seen that $k = K$ satisfies the requirements.

Third Case: $\dim l_t = 3$. If $l_t \cong \mathfrak{A}_i(\beta, 4\omega\beta^{-1}\pi^{2i}) \perp \langle\alpha\rangle$, $(\omega = 0$ or $\rho)$, is the canonical form for l_t, and if l_t is modular, we set

$$k \cong \begin{cases} \mathfrak{A}_{i-1}(\beta\pi^{-2}, 4\omega\beta^{-1}\pi^{2i}) \perp \langle\alpha\rangle & \text{if } \beta\pi^{-2} \in \mathfrak{g}L, \\ \mathfrak{A}_{i-1}(\beta, 4\omega\beta^{-1}\pi^{2i-2}) \perp \langle\alpha\rangle & \text{if } \beta\pi^{-2} \notin \mathfrak{g}L. \end{cases}$$

The reader can easily check that this is a proper choice of k; the only difficulty that arises is the case where $|\beta| = |2\pi^i|$, $\beta\pi^{-2} \notin \mathfrak{g}L$. But then the existence of K shows that $\omega = 0$, and the choice of k is legitimate. If l_t is almost modular, (so that α is not necessarily a norm generator of $\mathfrak{g}l_t$), the considerations are very similar to those of the modular case and are omitted; we merely point out that the existence of K shows that $\alpha\pi^{-2} \in \mathfrak{g}L$. This finishes the proof of Theorem 6.6.

THEOREM 6.7. *Let L be unimodular, and l another lattice with* $\dim l = \dim L \geqq 3$ *and* $\dim l_t = 2$ *where l_t is the last component in a Jordan splitting* $l = l_1 \perp \cdots \perp l_t$. *Suppose further that* $\mathfrak{n}L \subseteq \mathfrak{n}l^{\#\mathfrak{s}} = \mathfrak{m}$, *where* $\mathfrak{s} = \mathfrak{s}l_{t-1}{}^{\#}$. *Then $l \rightarrow L$ if and only if* $Fl \cong FL$, $\mathfrak{s}l \subseteq \mathfrak{s}L$, *and* $\mathfrak{g}l \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{m}}{}^{\#}$.

Proof. The necessity is obvious (note that $l \rightarrow L$ implies $L \rightarrow l_{\mathfrak{m}}{}^{\#}$ by 4.3). Therefore suppose that $Fl \cong FL$, $\mathfrak{s}l \subseteq \mathfrak{s}L$, $\mathfrak{g}l \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{m}}{}^{\#}$. Take a Jordan splitting $l = l_1 \perp \cdots \perp l_t$ and let $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$. We can suppose that $\mathfrak{n}L \supset 2\mathfrak{o}$ since otherwise $l \rightarrow L$ by 5.5 and 5.3.

First Case: $\mathfrak{n}l^{\#\mathfrak{s}} = \mathfrak{n}l^{\#}$. Then by definition of $l_{\mathfrak{m}}$, $l_{\mathfrak{m}}{}^{\#} = l^{\#}$. Suppose that $\mathfrak{n}l_{t-1}{}^{\#} \subset \mathfrak{n}l^{\#} = \mathfrak{n}l^{\#\mathfrak{s}}$. Then $l^{\#\mathfrak{s}} = l_{t-1}{}^{\#} \perp K$ for some lattice K with $\mathfrak{n}K = \mathfrak{n}l^{\#\mathfrak{s}}$ (cf. 4.2); $l_{t-1}{}^{\#}$ is the first component in a Jordan splitting of $l^{\#\mathfrak{s}}$. A direct application of Lemma 3.10 shows that another Jordan splitting can be chosen for $l^{\#\mathfrak{s}}$ in which the first component k has $\mathfrak{n}k = \mathfrak{n}l^{\#\mathfrak{s}}$. Since $k \subseteq l^{\#\mathfrak{s}}$, it splits $l^{\#}$ and it is clear that it can be chosen as the second component in a Jordan splitting of $l^{\#}$. Therefore we may assume that $l = l_1 \perp \cdots \perp l_t$ has the property that $\mathfrak{n}l_t{}^{*\#} = \mathfrak{n}l^{\#} = \mathfrak{n}l^{\#\mathfrak{s}}$.

Choose $c \in Q(l_t{}^{*\#})$ with $\mathfrak{n}l^{\#} = c\mathfrak{o}$ and write $l_t{}^{\#} \cong \mathcal{B}_{-i}(\alpha, \beta)$. Suppose that $\alpha \sim c$; then we can "reduce" α, using c, by the same method as in Lemma 3.11 (i.e., using the perfectness of the residue class field along with Lemma 3.10). A similar thing could be done if $\beta \sim c$. Let k be the lattice derived from $l_t{}^{\#}$ by one application of this technique. It is easily seen that $\mathfrak{n}k^{\#} = \mathfrak{n}l^{\#}$ (where $l^{\#} = k^{*} \perp k$), and therefore the reduction method can be applied again to k if necessary. It is clear that a sufficient number of applications of this method will yield a lattice of the form $\mathcal{A}_{-i}(\gamma, 4\omega\gamma^{-1}\pi^{-2i})$, $\omega = 0$ or ρ, where $|c| > |\gamma| \geqq |4\omega\gamma^{-1}\pi^{-2i}|$, $c \not\sim \gamma$. Let us assume that $l_t{}^{\#}$ is already in this form.

Since $\mathfrak{g}l \subseteq \mathfrak{g}$L, it is evident that $\pi^{2i}\gamma \in \mathfrak{g}L$. If $\pi^{2i-2}\gamma \in \mathfrak{g}L$, we put

$$J \cong l_t{}^{*} \perp \mathcal{B}_{i-1}(\pi^{2i-2}\gamma, 4\omega\gamma^{-1});$$

clearly $l \rightarrow J$, $\mathfrak{g}J \subseteq \mathfrak{g}L \subseteq gJ^{\#} = \mathfrak{g}l^{\#}$. If $\pi^{2i-2}\gamma \notin \mathfrak{g}L$, it is easly to see that $\pi^{2i}\gamma$ is a base generator of $\mathfrak{g}L$. Then the existence of a $\mathfrak{p}^{i-1}$-modular lattice on Fl_t with norm group in $\mathfrak{g}L$ (Lemma 5.2) shows that $|\pi^{2i}\gamma| > |4\omega\gamma^{-1}|$; in this case define

$$J \cong l_t{}^{*} \perp \mathcal{A}_{i-1}(\pi^{2i}\gamma, 4\omega\gamma^{-1}\pi^{-2});$$

again, $l \rightarrow J$, $\mathfrak{g}J \subseteq \mathfrak{g}L$. Clearly $b \in \gamma\mathfrak{p}^2 \subseteq \mathfrak{g}J^{\#}$. Now $\mathfrak{g}l_t{}^{*\#}$ contains an element with value $|a|$ since $c \sim a$ and $\mathfrak{n}L \subseteq \mathfrak{n}l^\# = \mathfrak{n}l_t^{*\#}$; we can suppose that it is of the form $a + d$ with $\mathfrak{d}(a^2 + ad) = ad\mathfrak{o}$. Then $\pi^{2j}(a + d) \in \mathfrak{g}L$, where $\mathfrak{s}l_{t-1} = \mathfrak{p}^j$, so that $\pi^{2j}d \in b\mathfrak{o} = \gamma\mathfrak{p}^{2i}$; therefore $d \in \gamma\mathfrak{p}^2$ and so $a \in \mathfrak{g}J^\#$. Thus again $\mathfrak{g}L \subseteq \mathfrak{g}J^\#$.

By induction (on $i - j$), we get a lattice K on Fl such that $l \rightarrow K$, $\mathfrak{g}K \subseteq \mathfrak{g}L \subseteq \mathfrak{g}K^\#$, $\mathfrak{s}K \subseteq \mathfrak{s}L$, and which has a Jordan form $K = K_1 \perp \cdots \perp K_{t-1}$ with $\dim K_{t-1} \geqq 3$. Therefore $K \rightarrow L$ by Theorem 6. 6, and so $l \rightarrow L$ as required.

Second Case: $\mathfrak{n}l^{\#\mathfrak{s}} \subset \mathfrak{n}l^\#$. Then for every Jordan splitting $l = l_1 \perp \cdots \perp l_t$, $\mathfrak{n}l^\# = \mathfrak{n}l_t^\#$; by the same method used in the First Case, we can choose the Jordan splitting so that $\mathfrak{n}l_{t-1}^\# = \mathfrak{n}l^{\#\mathfrak{s}}$.

1). $l_t \cong \mathcal{H}_i$. If $\mathfrak{n}l_{t-1}^\# \sim b$, choose γ so that $\gamma\mathfrak{o} = \pi\mathfrak{n}l_{t-1}^\#$. If $\mathfrak{n}l_{t-1}^\# \nsim b$, let γ be an element of maximal value in $2\mathfrak{p}^{-i} \cap b\mathfrak{p}^{-2j} \cap \mathfrak{p}^{-j}$ (where $\mathfrak{s}l_{t-1} = \mathfrak{p}^j$) such that $\gamma \sim b$. Then straightforward calculations show that $l_t^\# \cong \mathcal{A}_{-i}(\gamma,)$ and $l \rightarrow l_t^* \perp \mathcal{A}_j(\gamma\pi^{2j}, 0) \rightarrow L$ (using Theorem 6. 6).

2). $l_t \cong \mathcal{B}_i(\alpha, \beta) \ncong \mathcal{H}_i$, $\beta\pi^{-2} \notin \mathfrak{g}L$. Then it is easy to see that β is a base generator of $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$. By Lemma 5. 2, there is a $\mathfrak{p}^j$-modular lattice $P \cong \mathcal{B}_j(\pi^{2j-2i}\alpha', \beta')$ on Fl_t such that $\mathfrak{g}P \subseteq \mathfrak{g}L$. By replacing α' by $\alpha'\pi^{2k}$ and β' by $\beta'\pi^{-2k}$ if necessary, we may suppose that $|\alpha'| = |\alpha|$ and $|\beta'| = |\beta|$. But it is easy to see that $\mathfrak{g}l_t = \mathfrak{g}\mathcal{B}_i(\alpha', \beta')$, (express α and α' as elements of $a\mathfrak{o}^2 + b\mathfrak{o}$), and therefore $l_t \cong \mathcal{B}_i(\alpha', \beta') \rightarrow P$. Since β is a base generator of $\mathfrak{g}L$ and since $\mathfrak{g}P \subseteq \mathfrak{g}L$, there is a norm generator for $\mathfrak{g}P$ of the form $\xi^2 a$. Since $\mathfrak{n}P^\# = \mathfrak{n}l^\# \supset \mathfrak{n}L$, it follows that $\mathfrak{g}L \subseteq \mathfrak{g}P^\#$. Therefore $l \rightarrow l_t^* \perp P \rightarrow L$ by Theorem 6. 6.

3). $l_t \cong \mathcal{B}_i(\alpha, \beta) \ncong \mathcal{H}_i$, $\beta\pi^{-2} \in \mathfrak{g}L$. If we let $K \cong l_t^* \perp \mathcal{A}_{i-1}(\alpha, \beta\pi^{-2})$, then $\mathfrak{g}K \subseteq \mathfrak{g}L$ and $l \rightarrow K$. Since l_t is not a hyperbolic plane, the principle of domination shows that $l_{\mathfrak{m}}^\# \subseteq K^\#$. But $l^{\#\mathfrak{s}} \subset K^{\#\mathfrak{s}}$ so that $\mathfrak{n}l^{\#\mathfrak{s}} \subseteq \mathfrak{n}K^{\#\mathfrak{s}} = \mathfrak{m}'$, whence $l_{\mathfrak{m}}^\# \subseteq K_{\mathfrak{m}'}^\#$. Therefore $\mathfrak{g}L \subseteq \mathfrak{g}K_{\mathfrak{m}'}^\#$. The proof of 3) now follows by induction on $i - j$. Indeed if $i - j = 1$, then $l \rightarrow K \rightarrow L$ by Theorem 6. 6. Suppose $i - j > 1$. If $\mathfrak{n}K^{\#\mathfrak{s}} = \mathfrak{n}K^\#$, then $l \rightarrow K \rightarrow L$ by the First Case of this theorem. If $k \cong \mathcal{A}_{i-1}(\alpha, \beta\pi^{-2}) \cong \mathcal{H}_{i-1}$ then $l \rightarrow K \rightarrow L$ by 1), or else we have $k \cong \mathcal{B}_{i-1}(\alpha', \beta') \ncong \mathcal{H}_{i-1}$. If $\beta'\pi^{-2} \notin \mathfrak{g}L$, then $l \rightarrow K \rightarrow L$ by 2), while if $\beta'\pi^{-2} \in \mathfrak{g}L$, $l \rightarrow K \rightarrow L$ by the induction hypothesis.

LEMMA 6. 8. *Suppose that* $K \cong J \perp \mathcal{B}(\alpha, \beta)$, $K' \cong J' \perp \mathcal{B}(\alpha', \beta')$, *where* $\mathfrak{g}J = \mathfrak{g}J'$, $\mathfrak{n}J \subset \mathfrak{n}K = \alpha\mathfrak{o}$, $\mathfrak{n}J' \subset \mathfrak{n}K' = \alpha'\mathfrak{o} = \alpha\mathfrak{o}$, *and* $|\beta| = |\beta'|$. *Let a be a norm generator of* $\mathfrak{g}K$. *Then* $\mathfrak{g}K' = \mathfrak{g}K$ *if* $\mathfrak{d}(a\alpha') \subseteq \mathfrak{d}(a\alpha)$.

Proof. We can assume that $\alpha = a + \gamma$ with $\mathfrak{d}(a\alpha) = a\gamma\mathfrak{o}$, and $\alpha' = a + \gamma'$

with $\mathfrak{d}(a\alpha') = a\gamma'\mathfrak{o}$; thus $|\gamma'| \leqq |\gamma|$. Since it is clear that $\mathfrak{g}K' = \mathfrak{g}K$ if $\gamma \in \mathfrak{b} = 2\mathfrak{o} + \beta\mathfrak{o}$, we can suppose that $\gamma \notin \mathfrak{b}$.

Since $a \in \mathfrak{g}K$, we have $a = c + \epsilon^2(a + \gamma) + b'$ where $b' \in \mathfrak{b}$, $c \in \mathfrak{g}J$, and $\epsilon \in \mathfrak{u}$ since $\mathfrak{n}J \subset \mathfrak{n}K$. Write $c = \xi^2 a + \eta$ with $\mathfrak{d}(\xi^2 a^2 + a\eta) = a\eta\mathfrak{o}$. Then

$$a^2 \equiv (\epsilon + \xi)^2 a^2 + \epsilon^2 \gamma a + \eta a \bmod a\mathfrak{b}$$

and a comparison of the quadratic defects of both sides of this congruence shows that $|\gamma| = |\eta|$. Now if $\gamma' \in \mathfrak{b}$, $a \in \mathfrak{g}\mathfrak{B}(a + \gamma', \beta')$; suppose that $\gamma' \notin \mathfrak{b}$, so that $\gamma' \sim \gamma \nsim a$. Choose $\lambda \in \mathfrak{o}$ such that $|\lambda^2\eta + \gamma'| < |\gamma'|$. Then $\lambda^2 c + a + \gamma' \in \mathfrak{g}K'$ since $c \in \mathfrak{g}J'$, and so $(1 + \lambda\xi)^2 a + (\lambda^2\eta + \gamma')$ is a norm generator of $\mathfrak{g}K'$, (note that $\xi \in \mathfrak{p}$ since $\mathfrak{n}J \subset a\mathfrak{o}$). Thus there is a norm generator of the form $a + \gamma''$ for $\mathfrak{g}K'$, where $|\gamma''| < |\gamma'|$:

$$\mathfrak{g}K' = (a + \gamma'')\mathfrak{o}^2 + \beta''\mathfrak{o}.$$

But $a + \gamma' \in \mathfrak{g}K'$ implies that $\gamma' \in \beta''\mathfrak{o}$, and therefore $a \in \mathfrak{g}K'$

Now let b be a base generator of $\mathfrak{g}K$. If $b \in \mathfrak{b}$, then $b \in \mathfrak{g}K'$. If b is not in $\mathfrak{b}$ but in $\gamma\mathfrak{o}$, then $b \in \eta\mathfrak{o} \subset \mathfrak{g}K'$ since $\eta = (\xi^2 a + \eta) - \xi^2 a \in \mathfrak{g}K'$ by the first part. If $b \notin \gamma\mathfrak{o} + \mathfrak{b}$, then $b \in \mathfrak{g}K$ implies that

$$b \equiv \theta^2 a + \zeta + \delta^2 a \bmod \gamma\mathfrak{o} + \mathfrak{b}$$

where $\mathfrak{d}(\theta^2 a^2 + \zeta a) = \zeta a\mathfrak{o}$, $\theta^2 a + \zeta \in \mathfrak{g}J$. And much the same considerations as previously show that $|b| = |\zeta|$. Therefore $b \in \zeta\mathfrak{o} \subset \mathfrak{g}K'$.

We have shown thus far that the generators a and b are in $\mathfrak{g}K'$; therefore $\mathfrak{g}K \subseteq \mathfrak{g}K'$. But $|\gamma'| \leqq |\gamma| \leqq |b|$ implies that $\alpha' \in \mathfrak{g}K$, and similarly $\beta' \in \mathfrak{g}K$. Thus $\mathfrak{g}K' \subseteq \mathfrak{g}K$ and $\mathfrak{g}K' = \mathfrak{g}K$ as required.

THEOREM 6.9. *Let L be unimodular, and l another lattice with* $\dim l = \dim L \geqq 3$ *and* $\dim l_t = 2$ *where l_t is the last component in a Jordan splitting $l = l_1 \perp \cdots \perp l_t$. Suppose further that* $\mathfrak{n}l^{\#\mathfrak{s}} \subset \mathfrak{n}L = \mathfrak{m}$, *where* $\mathfrak{s} = \mathfrak{s}l_{t-1}{}^{\#}$, *and that a is a norm generator of* $\mathfrak{g}L$. *Then $l \to L$ if and only if* $Fl \cong FL$, $\mathfrak{s}l \subseteq \mathfrak{s}L$, $\mathfrak{g}l_{\mathfrak{m}} \subseteq \mathfrak{g}L \subseteq \mathfrak{g}\langle a \rangle + \mathfrak{g}l_{\mathfrak{m}}{}^{\#\mathfrak{s}} \subseteq \mathfrak{g}l_{\mathfrak{m}}{}^{\#}$, *and whenever* $\mathfrak{s} \subset \mathfrak{s}l_{\mathfrak{m}}{}^{\#}$, *then* $a \in Q(l_{\mathfrak{m}}{}^{\#})$.

Remarks. The condition $\mathfrak{n}l^{\#\mathfrak{s}} \subset \mathfrak{n}L$ will not hold if $\dim L$ is odd, in particular if $\dim L = 3$; indeed if $\dim L$ is odd, then so is $\dim l_t{}^{*}$, whence $\mathfrak{n}l^{\#\mathfrak{s}} \supseteq \mathfrak{o}$. In Lemma 6.10, we shall show how to tell if $a \in Q(l_{\mathfrak{m}}{}^{\#})$.

Proof of Necessity. If $l \to L$, then by 4.3, $l_{\mathfrak{m}} \to L$, whence

$$\mathfrak{g}l_{\mathfrak{m}} \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{m}}{}^{\#} \text{ and } a \in \mathfrak{g}L = Q(L) \subseteq Q(l_{\mathfrak{m}}{}^{\#})$$

—the fact that $\mathfrak{g}L = Q(L)$ when $\dim L \geqq 4$ follows from elementary considerations or from the Third Main Theorem, (iii), in § 7; we can use this theorem as a reference since it is perfectly clear from its proof that (iii) is independent of the present theorem. If $\mathfrak{s}l_{\mathfrak{m}}{}^{\#} \subseteq \mathfrak{s}$, then $l_{\mathfrak{m}}{}^{\#\mathfrak{s}} = l_{\mathfrak{m}}{}^{\#}$, and so $\mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{m}}{}^{\#} = \mathfrak{g}\langle a\rangle + \mathfrak{g}l_{\mathfrak{m}}{}^{\#\mathfrak{s}}$. Suppose $\mathfrak{s} \subset \mathfrak{s}l_{\mathfrak{m}}{}^{\#} = \mathfrak{p}^{-m}$; by Lemma 4. 4, $l_{\mathfrak{m}} = J \perp K$ where K is binary $\mathfrak{p}^{m}$-modular (and $J^{\#} \subset l^{\#\mathfrak{s}}$).

Suppose first that $K \cong \mathcal{H}_m$. By Theorem 5. 1, there is an $\mathfrak{s}$-modular lattice $k^{\#}$ on FK such that $L \to J^{\#} \perp k^{\#}$. Since $\mathfrak{n}J^{\#} \subset \mathfrak{n}L$, we have $\mathfrak{n}k^{\#} \supseteq \mathfrak{n}L$; explicitly, $k^{\#} = \mathfrak{o}x + \mathfrak{o}y \cong \mathcal{B}_{-j}(a', 0)$ where $\mathfrak{s} = \mathfrak{p}^{-j}$. If $\mathfrak{n}k^{\#} \supset \mathfrak{n}L$, then by domination, $L \to J^{\#} \perp (\mathfrak{o}\pi x + \mathfrak{o}\pi^{-1}y)$; it is therefore clear that we may assume that $| a' | = | a |$.

Suppose now that $K \ncong \mathcal{H}_m$, say $K^{\#} \cong \mathcal{B}_{-m}(\alpha, \beta)$. Then by Lemma 4. 5, it is easily seen that $\mathfrak{n}l_{\mathfrak{m}}{}^{\#} = \mathfrak{n}L$, i. e., $| \alpha | = | a |$. But $a \in Q(l_{\mathfrak{m}}{}^{\#})$, and therefore it is easy to see that, (by choosing another splitting $J' \perp K'$ if necessary), we can assume that $\alpha = a$. Again by Theorem 5. 1, there is an $\mathfrak{s}$-modular lattice $k^{\#}$ on FK such that $L \to J^{\#} \perp k^{\#} \subset l_{\mathfrak{m}}{}^{\#}$; it is clear that $\mathfrak{n}L = \mathfrak{n}k^{\#}$. Let us write $k^{\#} \cong \mathcal{B}_{-j}(a', \beta')$ where $| a' | = | a |$. A comparison of $\mathfrak{d}(-dK)$ and $\mathfrak{d}(-dk)$ shows that $| \beta' | = | \pi^{2m-2j}\beta |$.

Now combine the two cases by setting $\beta = \beta' = 0$ if $K \cong \mathcal{H}_m$. We have $\mathfrak{g}L \subseteq \mathfrak{g}(J \perp k)^{\#}$ and a is a norm generator of $\mathfrak{g}(J \perp k)^{\#}$. Moreover it is easy to see that

$$\mathfrak{g}\langle a\rangle + \mathfrak{g}l_{\mathfrak{m}}{}^{\#\mathfrak{s}} = \mathfrak{g}(J^{\#} \perp \mathcal{B}_{-j}(a, \pi^{2m-2j}\beta));$$

but this latter norm group is equal to $\mathfrak{g}(J \perp k)^{\#}$ by Lemma 6. 8, and therefore $\mathfrak{g}L \subseteq \mathfrak{g}\langle a\rangle + \mathfrak{g}l_{\mathfrak{m}}{}^{\#\mathfrak{s}}$. Since the containment $\mathfrak{g}\langle a\rangle + \mathfrak{g}l_{\mathfrak{m}}{}^{\#\mathfrak{s}} \subseteq \mathfrak{g}l_{\mathfrak{m}}{}^{\#}$ is obvious, the proof of the necessity of Theorem 6. 9 is finished.

Proof of the Sufficiency. It suffices to prove that $l_{\mathfrak{m}} \to L$. Split $l_{\mathfrak{m}} = J \perp K$ as in the proof of the necessity; thus K is a binary $\mathfrak{p}^{m}$-modular lattice.

First Case: $K \cong \mathcal{H}_m$. Then

$$\mathfrak{s} \subset \mathfrak{s}K^{\#},\ K^{\#} \cong \mathcal{A}_{-m}(a, 0),\ \text{and}\ l_{\mathfrak{m}} \to M \cong J \perp \mathcal{A}_{j}(\pi^{2j}a, 0) \to L$$

by Theorem 6. 6, since $\mathfrak{g}L \subseteq \mathfrak{g}M^{\#} = \mathfrak{g}\langle a\rangle + \mathfrak{g}l_{\mathfrak{m}}{}^{\#\mathfrak{s}}$.

Second Case: $\mathfrak{s} \subset \mathfrak{s}K^{\#}$, $K \ncong \mathcal{H}_m$. Since $a \in Q(l_{\mathfrak{m}}{}^{\#})$, and $\mathfrak{n}l_{\mathfrak{m}}{}^{\#} = \mathfrak{n}L$ by Lemma 4. 5, we can suppose that $K^{\#} \cong \mathcal{B}_{-m}(a, \beta)$. Put $l' \cong J \perp \mathcal{B}_{j}(\pi^{2j}a, \pi^{2m}\beta)$. Clearly $l_{\mathfrak{m}} \to l'$, $\mathfrak{g}l' \subseteq \mathfrak{g}L \subseteq \mathfrak{g}\langle a\rangle + \mathfrak{g}l_{\mathfrak{m}}{}^{\#\mathfrak{s}} = \mathfrak{g}l'^{\#}$. Therefore $l' \to L$ by Theorem 6. 6, and so $l \to L$.

Third Case: $\mathfrak{s}K^{\#} \subseteq \mathfrak{s}$. We can assume that the splitting $l_{\mathfrak{m}} = J \perp K$ has been chosen so that $J = l_t{}^*$ for some Jordan splitting $l = l_1 \perp \cdots \perp l_t$; thus $l_t \subset K$. By Theorem 5.1, there is a $\mathfrak{p}^j$-modular lattice P on FK such that $l_t \subset P \subseteq K$. Then

$$\mathfrak{g}(l_t{}^* \perp P) \subseteq \mathfrak{g}l_{\mathfrak{m}} \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{m}}{}^\# \subseteq \mathfrak{g}(l_t{}^* \perp P)^\#$$

and so $l \to l_t{}^* \perp P \to L$. This finishes the proof of Theorem 6.9.

The purpose of the next lemma is to enable one to determine effectively whether or not the condition $a \in Q(l_{\mathfrak{m}}{}^\#)$ of Theorem 6.9 is fulfilled. Note that the defect sum in the statement of 6.10 is equal to $\mathfrak{d}(a\alpha) + \mathfrak{d}(a\beta)$ where $\mathfrak{g}L_1{}^* = \alpha\mathfrak{o}^2 + \beta\mathfrak{o}$.

LEMMA 6.10. *Let L be a lattice with a splitting $L = K_1 \perp K_1{}^*$ with the following properties: K_1 is binary unimodular, $\mathfrak{s} = \mathfrak{s}K_1{}^* \subset \mathfrak{s}K_1$, $\mathfrak{n}L^{\mathfrak{s}} \subset \mathfrak{n}L$, $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o} \subseteq \mathfrak{s}$. Then $a \in Q(L)$ if and only if $a \in \mathfrak{g}L$ and, if $L = L_1 \perp L_1{}^*$ is a splitting of the above sort, then at least one of the following holds:*

$$a \in Q(FL_1), \text{ or } ab \notin 2\mathfrak{o}, \text{ or } 4ab^{-1} \in \sum \mathfrak{d}(ac),$$

where the sum runs over $c \in \mathfrak{g}L_1{}^*$.

Proof of Necessity. Suppose $a \notin Q(FL_1)$ and $ab \in 2\mathfrak{o}$. Write

$$L_1 \cong \mathcal{B}(a + \gamma, \beta),$$

where $\mathfrak{d}(a^2 + a\gamma) = a\gamma\mathfrak{o}$ is smallest for all such expressions for L_1. Since $a \in Q(L)$, we have

$$a = \epsilon^2(a + \gamma) + 2\epsilon\lambda + \lambda^2\beta + \xi^2 a + \eta$$

with $\epsilon \in \mathfrak{u}$, $\lambda \in \mathfrak{o}$, $\xi^2 a + \eta \in \mathfrak{g}L_1{}^*$, $\mathfrak{d}(\xi^2 a^2 + a\eta) = a\eta\mathfrak{o}$; note that $\xi \in \mathfrak{p}$. Suppose that $\beta \in \gamma\mathfrak{p}$. Since γ is minimal, $a \notin Q(FL_1)$, and Lemma 3.4, we have $\gamma \notin 2\mathfrak{o}$. Also

$$a^2 \equiv (\epsilon^2(a + \gamma) + \xi^2 a + \eta)a \equiv ((\epsilon + \xi)^2 a + \epsilon^2\gamma + \eta)a \bmod a\gamma\mathfrak{p}.$$

A comparison of the quadratic defects of the two extreme members of this congruence shows that $|\gamma| = |\eta|$ whence $4ab^{-1} \in 4a\beta^{-1}\mathfrak{o} \subseteq a\gamma\mathfrak{o} = a\eta\mathfrak{o}$ as required; the second containment follows by virtue of the fact that $a \notin Q(FL_1)$ implies that $\mathfrak{d}(-dL_1) \sim \pi$, and Lemma 3.7.

Suppose that $\beta \notin \gamma\mathfrak{o}$. By Lemma 3.7 again, we can assume that $|\gamma| = |4\beta^{-1}|$. Then the principle of domination and Remark 3.8 show that

$$a^2 \equiv ((\epsilon + \xi)^2 a + \zeta + \eta)a \bmod 2a^2\mathfrak{p}$$

with $|\zeta| \geqq |\gamma|$, $\zeta \sim \gamma$. If $ab \in 2\mathfrak{o}$, $|\zeta| \geqq |4\beta^{-1}| \geqq |4b^{-1}| \geqq |2a|$; then comparing defects in the congruence leads to $|\zeta| = |\eta|$ and the necessity follows as above.

Proof of Sufficiency. First let us show that L_1 can be chosen so that $\mathfrak{g}L_1 = \mathfrak{g}L$. Since $a \in \mathfrak{g}L$ and $\mathfrak{n}L^{\mathfrak{s}} \subset \mathfrak{n}L$, L can be expressed as

$$L \cong \mathcal{B}(a+\gamma, \beta) \perp K$$

where $\gamma \in 2\mathfrak{o}$, and $\mathfrak{d}(a^2 + a\gamma) = a\gamma\mathfrak{o}$. If $\mathfrak{g}\mathcal{B}(a+\gamma, \beta) \subset \mathfrak{g}L$, then (cf. formula for $\mathfrak{g}_1 + \mathfrak{g}_2$ before Lemma 3.4), we can find $c \in Q(K)$ such that

$$(a+\gamma)^{-1}\mathfrak{d}(c(a+\gamma)) = b\mathfrak{o}.$$

The lattice $L_1 \cong \mathcal{A}(a+\gamma, \beta + c)$ splits L by 3.10, and $\mathfrak{g}L_1 = \mathfrak{g}L$; write $L_1 \cong \mathcal{B}(a+\gamma, b')$. Note that we can assume $\mathfrak{n}L \supset 2\mathfrak{o}$ since otherwise $a \in Q(L_1)$ easily.

First suppose that $b' \in 2\mathfrak{p}$; then since $\gamma \in 2\mathfrak{o}$, $L_1 \cong \mathcal{B}(a, b'')$ by Lemma 3.4. Thus we can assume that $b' \notin 2\mathfrak{p}$, and so b' $(= b$, say) is a base generator of $\mathfrak{g}L$.

Next suppose that $a \in Q(FL_1)$. Then clearly $L_1 \cong \mathcal{B}(a, b'')$. Therefore we assume that $a \notin Q(FL_1)$, whence γ can be assumed to have value $|\gamma| = |4\delta^{-1}a| = |4b^{-1}|$.

If $2b^{-1} \in \mathfrak{s} = \mathfrak{p}^i$, then $\gamma \in 2\mathfrak{s}$, and so a is a norm generator of

$$L' \cong \mathcal{B}_i(a+\gamma, \pi^{2i}b) \perp L_1^{*} \rightarrow L.$$

The conditions $a \in \mathfrak{s}$ and $\mathfrak{n}L^{\mathfrak{s}} \subset \mathfrak{n}L$ show that the first component L_1' in a Jordan splitting of L' has $\dim L_1' \geqq 4$; using the method employed at the beginning of the proof of Theorem 6.6, we can find another Jordan splitting of L' in which the first component K_1 has $\mathfrak{g}K_1 = \mathfrak{g}L'$. By the Third Main Theorem (iii) of § 7, (which is independent of this lemma), $\mathfrak{g}K_1 = Q(K_1)$, and so $a \in Q(K_1) \subseteq Q(L)$.

Finally suppose $2b^{-1} \notin \mathfrak{s}$; then $ab \in 2\mathfrak{o}$, and so we can find $\xi^2 a + \eta \in \mathfrak{g}L_1^*$ such that $4ab^{-1} \in a\eta\mathfrak{o} = \mathfrak{d}(\xi^2 a^2 + a\eta)$. Now L_1^* represents $\xi^2 a + \eta \bmod 2\mathfrak{s}$; but $|2\mathfrak{s}| < |4b^{-1}| \leqq |\eta|$, so (by changing η to η' of the same value, if necessary), we can assume that $\xi^2 a + \eta \in Q(L_1^*)$. Choose $\lambda \in \mathfrak{o}$ such that $|\lambda^2\eta + \gamma| < |\gamma|$. The lattice $J \cong \mathcal{B}(a + \gamma + \lambda^2(\xi^2 a + \eta), b)$ splits L. It is easily checked, by Lemma 3.7, that $J \cong \mathcal{B}(a, b')$. This finishes the proof of Lemma 6.10.

THEOREM 6.11. *Let L be unimodular,* $\dim L \geqq 3$, *and let l be another lattice of the same dimension with* $\dim l_t = 1$ *where $l = l_1 \perp \cdots \perp l_t$ is a*

Jordan splitting of l. *Define* $\mathfrak{a} = \mathfrak{s}l_{t-1}{}^{\#}$, $\mathfrak{s} = \mathfrak{s}l_{t-2}{}^{\#}$ *and* $\mathfrak{m} = \mathfrak{n}L + \mathfrak{n}l^{\#\mathfrak{s}}$ *if* $t > 2$. *Let* a *be a norm generator of* $\mathfrak{g}L$. *Then necessary and sufficient conditions for* $l \to L$ *are* $Fl \cong FL$, $\mathfrak{s}l \subseteq \mathfrak{s}L$, *and:*

(i) $\mathfrak{g}l \subseteq \mathfrak{g}L$, *if* $\dim l_{t-1} = 1$ *and* $dl_{t-1} \not\sim dl_t$,

(ii) $\mathfrak{g}l_{\mathfrak{m}} \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{m}}{}^{\#}$,

$$\mathfrak{g}L \subseteq \mathfrak{g}\langle a\rangle + \mathfrak{g}l_{\mathfrak{m}}{}^{\#\mathfrak{s}} \qquad \textit{when } \mathfrak{n}l^{\#\mathfrak{s}} \subset \mathfrak{m} \tag{2}$$

$$a \in Q(l_{\mathfrak{m}}{}^{\#}) \qquad \textit{when } \mathfrak{n}l^{\#\mathfrak{s}} \subset \mathfrak{m} \textit{ and } \mathfrak{s} \subset \mathfrak{s}l_{\mathfrak{m}}{}^{\#} \tag{3}$$

if $\dim l_{t-1} = 1$ *and* $dl_{t-1} \sim dl_t$.

(iii) $\mathfrak{g}l_{\mathfrak{a}} \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{a}}{}^{\#}$, *if* $\dim l_{t-1} \geqq 2$.

Proof of Necessity. (i) is trivial. So are the first condition of (ii), and (iii), by 4.3. So suppose that $\dim l_{t-1} = 1$, $dl_{t-1} \sim dl_t$. Since $\mathfrak{m} \subset \mathfrak{a}$, it is easy to see (directly from the definitions) that $(l_{\mathfrak{a}})_{\mathfrak{m}} = l_{\mathfrak{m}}$ and that $\mathfrak{n}l_{\mathfrak{a}}{}^{\#\mathfrak{s}} = \mathfrak{n}l^{\#\mathfrak{s}}$. Moreover the last component in a Jordan splitting of $l_{\mathfrak{a}}$ is binary modular (scale $\mathfrak{s}l_{t-1}$). The conditions (2) and (3) now follow from Theorem 6.9 applied to $l_{\mathfrak{a}}$ and L, since $l \to L$ if and only if $l_{\mathfrak{a}} \to L$.

Proof of Sufficiency. Consider (i) first. By Lemma 5.2, there is an almost $\mathfrak{s}^{-1}$-modular lattice K on $F(l_{t-1} \perp l_t)$ with $\mathfrak{g}K = \mathfrak{s}^{-1} \subseteq \mathfrak{g}L$. But then Theorem 6.6 applies to $l_1 \perp \cdots \perp l_{t-2} \perp K = M$. It is easy to see that $\mathfrak{g}M \subseteq \mathfrak{g}L \subseteq \mathfrak{g}M_{\mathfrak{s}}{}^{\#} = \mathfrak{s}$ and therefore $M \to L$. But K is $\mathfrak{s}^{-1}$-maximal (cf. proof of Lemma 6.2) whence $l_{t-1} \perp l_t \to K$. Therefore $l \to M \to L$ as required.

Next (iii). Put $\mathfrak{b} = \mathfrak{a}$ if $\mathfrak{s}l_t \sim \mathfrak{s}l_{t-1}$, $\mathfrak{b} = \pi^{-1}\mathfrak{a}$ otherwise. Then $l_{\mathfrak{b}} \subseteq l_{\mathfrak{a}}$ and $l_{\mathfrak{b}}{}^{\#\mathfrak{a}} = l_{\mathfrak{a}}{}^{\#}$ so that $\mathfrak{g}l_{\mathfrak{b}} \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{b}}{}^{\#\mathfrak{a}}$. Therefore $l \to l_{\mathfrak{b}} \to L$ by Theorem 6.6.

Finally (ii). As in the proof of the necessity, $(l_{\mathfrak{a}})_{\mathfrak{m}} = l_{\mathfrak{m}}$ and $\mathfrak{n}l_{\mathfrak{a}}{}^{\#\mathfrak{s}} = \mathfrak{n}l^{\#\mathfrak{s}}$. Since the dimension of the last component in a Jordan splitting of $l_{\mathfrak{a}}$ is binary modular, one can apply either Theorem 6.7 (if $nl_{\mathfrak{a}}{}^{\#\mathfrak{s}} = \mathfrak{m}$) or Theorem 6.9 (if $\mathfrak{n}l_{\mathfrak{a}}{}^{\#\mathfrak{s}} \subset \mathfrak{m}$) to show that $l_{\mathfrak{a}} \to L$. Then $l \to l_{\mathfrak{a}} \to L$ as required.

THEOREM 6.12. *Let* L *be a binary unimodular lattice,* l *another lattice with* $\alpha \in Q(l)$ *a norm generator of* $\mathfrak{g}l$. *Then* $l \to L$ *if and only if* $Fl \to FL$, $\alpha \in Q(L)$, *and* $\mathfrak{n}L \subseteq \mathfrak{n}l^{\#}$ *if* l *is binary modular.*

Proof. The necessity is obvious. Write $L \cong \mathfrak{B}(a, c)$, and consider first the case where l is binary modular. If $ac \sim \pi$ and $\mathfrak{n}l \subseteq c\mathfrak{o}$, then $l \to L$ since L contains a $c\mathfrak{p}^{-1}$-maximal lattice by Lemma 3.6. If $ac \not\sim \pi$ and $\mathfrak{n}l \subseteq 4a^{-1}\mathfrak{o}$, then $l \to L$ since L contains a $4a^{-1}\mathfrak{o}$-maximal lattice by Lemma 5.4.

The elimination of the two above cases and the fact that $\alpha \in Q(L)$ means that

$$\alpha = \xi^2 a + 2\xi\eta + \eta^2 c,\ \xi \text{ and } \eta \text{ in } \mathfrak{o},$$

where $|\alpha| = |\xi^2 a| > \max\{|2\xi|, |c|\}$; this involves a straightforward computation which is omitted. Since $\mathfrak{n}L \subseteq \mathfrak{n}l^{\#}$, we have $|\pi^i| \leqq |\xi|$. Now.

$$L \leftarrow \begin{bmatrix} \xi^2 a & \xi \\ \xi & c \end{bmatrix} \cong K, \text{ say.}$$

Clearly $\alpha \in Q(K)$; in fact α must be a norm generator of $\mathfrak{g}K$ and therefore $K \cong \mathcal{B}_n(\alpha, \delta\alpha^{-1}\pi^{2n})$ where $\mathfrak{p}^n = \xi\mathfrak{o}$. But then, $K \leftarrow \mathcal{B}_i(\alpha, \delta\alpha^{-1}\pi^{2i}) \cong l$, and so $l \rightarrow L$ as required.

Next if l is binary but not modular, we have $l \rightarrow L$ if and only if $l_{\mathfrak{a}} \rightarrow L$ where $\mathfrak{a} = \alpha^{-1}\mathfrak{o}$, (cf. 4.3). But it is easy to see that $l_{\mathfrak{a}}$ is a binary $\alpha\mathfrak{o}$-modular lattice, and that α is in $Q(l_{\mathfrak{a}})$ and is a norm generator of $\mathfrak{g}l_{\mathfrak{a}}$. Therefore $l_{\mathfrak{a}} \rightarrow L$ by the first part, and so $l \rightarrow L$.

If $\dim l = 1$, the result is trivial since $l \cong \langle\alpha\rangle$.

We now want to show how to determine if $\alpha \in Q(L)$ in the situation of Theorem 6.12. This is particularly desirable since as the theorem stands, the conditions given for $\dim l = 1$ are tautological. Lemma 6.13 is of wider interest though, since it enables one to determine $Q(L)$ for a binary modular lattice L, (cf. Theorem 7.4).

LEMMA 6.13. *Let L be a binary unimodular lattice and let $\alpha \in Q(L)$ be a norm generator of $\mathfrak{g}L$. If $\alpha \in \mathfrak{n}L$, then $\alpha \in Q(L)$ if and only if $\alpha \in Q(FL)$ and*

$$(4) \qquad (\mathfrak{d}(a\alpha))^2 \subseteq 4\alpha\mathfrak{n}L + (\mathfrak{d}(-dL))^2.$$

Proof of Necessity. Since $\alpha \in Q(L)$, we have

$$(5) \qquad \alpha = \xi^2 a + 2\xi\eta + \eta^2 c,\ \xi \text{ and } \eta \in \mathfrak{o},$$

where $L \cong \mathcal{B}(a, c)$. Suppose $\mathfrak{d}(a\alpha) \supset \mathfrak{d}(-dL) = ac\mathfrak{o}$. If $\xi^2 a \in \alpha\mathfrak{o}$, then upon multiplying (5) by a, we see that $(\mathfrak{d}(a\alpha))^2 \subseteq (2\xi\eta a)^2\mathfrak{o} \subseteq 4\alpha\mathfrak{n}L$ as required. If $\xi^2 a \notin \alpha\mathfrak{o}$, then $c = 0$ (since otherwise the principle of domination leads to a contradiction), and it is easy to see that $|\xi^2 a| = |2\xi\eta| \leqq |2\xi|$. Thus $|\xi a| \leqq |2|$ and $|a\alpha| \leqq |4|$, whence $(\mathfrak{d}(a\alpha))^2 \subseteq (a\alpha)^2\mathfrak{o} \subseteq 4\alpha\mathfrak{n}L$ as required.

Proof of Sufficiency. If $a \not\sim \alpha$, then

$$(\mathfrak{d}(a\alpha))^2 = (a\alpha)^2\mathfrak{o} \subseteq 4a\alpha\mathfrak{o} + (ac)^2\mathfrak{o}$$

If $(a\alpha)^2 \in 4a\alpha\mathfrak{o}$, then $\alpha \in 4a^{-1}\mathfrak{o}$; if $a\alpha \in ac\mathfrak{o}$, then $\alpha \in c\mathfrak{o}$. In either case L contains an $\alpha\mathfrak{o}$-maximal lattice (by 5.4 or 3.6) and so $\alpha \in Q(L)$ by Lemma 5.3.

We can now suppose that $a \sim \alpha$, and also $\alpha \notin c\mathfrak{o} + 4a^{-1}\mathfrak{o}$. Since it suffices to show that $\alpha\epsilon^2 \in Q(L)$ for any unit ϵ, we can assume that $\alpha = \pi^{2i}a + \gamma$ where $\mathfrak{d}(a\alpha) = a\gamma\mathfrak{o} \subset a^2\mathfrak{p}^{2i} = a\alpha\mathfrak{o}$. The condition (4) is equivalent to $\gamma \in 2\mathfrak{p}^i + c\mathfrak{o}$; thus γ is in the norm group of the lattice $L' \cong \mathscr{B}_i(\pi^{2i}a, c)$—note that $\mathfrak{n}L' = a\mathfrak{p}^{2i}$ since $\alpha \notin c\mathfrak{o} + 4a^{-1}\mathfrak{o}$. By O'Meara's theorem on modular lattices $L' \cong \mathscr{B}_i(\alpha, c')$ for suitable c', and therefore $\alpha \in Q(L')$. But $i \geqq 0$ since $\alpha \in \mathfrak{n}L$, so that $L' \rightarrow L$ and therefore $\alpha \in Q(L)$ as required.

Remark. The following theorem gives a complete solution for $l \rightarrow L$ when $\dim l = \dim L$ and L is unimodular. For $\dim L \geqq 3$, it is essentially just a restatement of results given earlier in this section; on the other hand, the conditions given for $\dim L = 2$ are not those of Theorem 6.12, but present an alternative solution.

It is convenient to introduce the notation

$$\underset{1 \leqq j \leqq n}{\perp} K_j = K_1 \perp K_2 \perp \cdots \perp K_n.$$

FIRST MAIN THEOREM. *Let L be a unimodular lattice and let l be another lattice of the same dimension. If $\dim L \geqq 3$, let $\mathfrak{s} = \mathfrak{s}l_r^{\#}$ where $l = l_1 \perp \cdots \perp l_t$ is a Jordan splitting and r is the largest integer such that* $\dim (\underset{r \leqq j \leqq t}{\perp} l_j) \geqq 3$. *Put $\mathfrak{m} = \mathfrak{n}L + \mathfrak{n}l^{\#\mathfrak{s}}$ if $\dim L \geqq 3$, $\mathfrak{m} = \mathfrak{n}L$ if $\dim L \leqq 2$. Let a be a norm generator of $\mathfrak{g}L$. Then necessary and sufficient conditions for $l \rightarrow L$ are: $Fl \cong FL$, $\mathfrak{s}l \subseteq \mathfrak{s}L$, $\mathfrak{g}l_{\mathfrak{m}} \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{m}}^{\#}$ and*

$$\mathfrak{g}L \subseteq \mathfrak{g}\langle a\rangle + \mathfrak{g}l_{\mathfrak{m}}^{\#\mathfrak{s}} \textit{ if } \dim L \geqq 3 \textit{ and } \mathfrak{n}l^{\#\mathfrak{s}} \subset \mathfrak{m} \tag{2}$$

$$a \in Q(l_{\mathfrak{m}}^{\#}) \textit{ if } \dim \mathrm{L} \geqq 3,\ \mathfrak{n}l^{\#\mathfrak{s}} \subset \mathfrak{m} \textit{ and } \mathfrak{s} \subset \mathfrak{s}l_{\mathfrak{m}}^{\#}. \tag{3}$$

Remarks. The conditions under which (3) is non-empty are such that Lemma 6.10 can be used to decide whether or not $a \in Q(l_{\mathfrak{m}}^{\#})$.

The ideal $\mathfrak{m}$ can be defined simultaneously for all values of $\dim L$ by using the "partial Jordan splittings" introduced in the Concluding Remarks at the end of the paper. Namely put $k = l^{\#}$ and let $\mathfrak{s} = \bigcap_j \mathfrak{p}^j$ where j runs over all integers such that $\dim \mathfrak{k}_{j-1} \leqq 2$ (thus $\mathfrak{s}$ is the same ideal defined in the theorem if $\dim L \leqq 3$, while $\mathfrak{s} = 0$ if $\dim L \leqq 2$); then $\mathfrak{m} = \mathfrak{n}L + \mathfrak{n}l^{\#\mathfrak{s}}$ since $l^{\#\mathfrak{s}} = 0$ if $\mathfrak{s} = 0$.

Proof of Necessity. The conditions $Fl \cong FL$, $\mathfrak{s}l \subseteq \mathfrak{s}L$, and $\mathfrak{g}l_{\mathfrak{m}} \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{m}}^{\#}$ are obvious. Suppose $\dim L \geqq 3$ and $\mathfrak{n}l^{\#\mathfrak{s}} \subset \mathfrak{m}$. Then $\dim l_t \leqq 2$ since

$l^{\#\mathfrak{s}} = l^{\#}$ if $\dim l_t \geqq 3$. If $\dim l_t = 2$, (resp. if $\dim l_t = \dim l_{t-1} = 1$ and $dl_{t-1} \sim dl_t$), then (2) and (3) follow from Theorem 6.9 (resp. Theorem 6.11). If $\dim l_t = \dim l_{t-1} = 1$ and $dl_{t-1} \not\sim dl_t$, or if $\dim l_t = 1$ and $\dim l_{t-1} \geqq 2$, then a straightforward application of the principle of domination shows that $\mathfrak{s} = \mathfrak{s}l_{\mathfrak{m}}{}^{\#}$ so that (3) is empty. But $\mathfrak{s} = \mathfrak{s}l_{\mathfrak{m}}{}^{\#}$ implies that $l_{\mathfrak{m}}{}^{\#\mathfrak{s}} = l_{\mathfrak{m}}{}^{\#}$ and so (2) is fulfilled.

Proof of Sufficiency. If $\dim L = 1$, then $Fl \cong FL$ and $\mathfrak{s}l \subseteq \mathfrak{s}L$ clearly suffice. Suppose $\dim L \geqq 3$. If $\dim l_t \geqq 3$, then $l \to L$ by Theorem 6.6 since $l \subseteq l_{\mathfrak{m}}$ implies $\mathfrak{g}l \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l^{\#}$. If $\dim l_t = 2$, then $l \to L$ by either Theorem 6.7 or Theorem 6.9. Suppose $\dim l_t = 1$. If $\dim l_{t-1} = 1$, then the theorem follows directly from Theorem 6.11, (i) and (ii); note that in (ii), the ideals $\mathfrak{s}$ and $\mathfrak{m}$ are the same as defined here. If $\dim l_{t-1} \geqq 2$ (and $\dim d_t = 1$), then $\mathfrak{a} = \mathfrak{s}l_{t-1}{}^{\#} \supseteq \mathfrak{m}$ whence $l_{\mathfrak{a}} \subseteq l_{\mathfrak{m}}$. Thus $\mathfrak{g}l_{\mathfrak{a}} \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{a}}{}^{\#}$ and so $l \to L$ by Theorem 6.11 (iii), as required.

Now suppose that $\dim L = 2$. If l is not modular, it is easy to see by the principle of domination that $(l_{\mathfrak{a}})_{\mathfrak{m}} = l_{\mathfrak{m}}$ where $\mathfrak{a} = (\mathfrak{n}l)^{-1}$. Since $l_{\mathfrak{a}}$ is modular and $l \to l_{\mathfrak{a}}$, we may therefore assume that l is modular. By Lemmas 4.4 and 4.5 $l_{\mathfrak{m}}$ is either a hyperbolic plane $\mathcal{H}_i$ or a binary modular lattice with $\mathfrak{n}l_{\mathfrak{m}}{}^{\#} = \mathfrak{m} = \mathfrak{n}l$ (since $\mathfrak{n}L \subseteq \mathfrak{n}l_{\mathfrak{m}}{}^{\#}$). If $l_{\mathfrak{m}} \cong \mathcal{H}_i$, then $l_{\mathfrak{m}}{}^{\#} \cong \mathcal{H}_{-i} \cong \mathcal{A}_{-i}(a, 0)$ and so $l \to l_{\mathfrak{m}} \cong \mathcal{A}_i(\pi^{2i}a, 0) \to \mathcal{A}(a, 0) \cong L$, since $\mathfrak{g}l_{\mathfrak{m}} \subseteq \mathfrak{g}l_{\mathfrak{m}}{}^{\#}$ implies that $i \geqq 0$.

On the other hand if $\mathfrak{n}l_{\mathfrak{m}}{}^{\#} = \mathfrak{n}L$, then $l_{\mathfrak{m}}{}^{\#} \cong \mathcal{B}_{-i}(\alpha, \delta\alpha^{-1}\pi^{-2i})$ where $dL \cong -1 + \delta$, $\mathfrak{d}(1 - \delta) = \delta\mathfrak{o}$, and α is a norm generator of $\mathfrak{g}L$ contained in $Q(L)$; since $\mathfrak{n}l_{\mathfrak{m}}{}^{\#} = \mathfrak{n}L$ and $\mathfrak{g}L \subseteq \mathfrak{g}l_{\mathfrak{m}}{}^{\#}$, α is also a norm generator of $\mathfrak{g}l_{\mathfrak{m}}{}^{\#}$. Thus $l \to l_{\mathfrak{m}} \cong \mathcal{B}_i(\pi^{2i}\alpha, \delta\alpha^{-1}) \to \mathcal{B}(\alpha, \delta\alpha^{-1}) \cong L$ as required.

The next theorem is largely contained in the First Main Theorem, and is somewhat weaker since the condition $Q(l) \subseteq Q(L) \subseteq Q(l^{\#})$ is stronger in general than $\mathfrak{g}l \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l^{\#}$. Nevertheless it was considered worthwhile to single the result out since the modular case is fundamental for the representation theory, as it is for the equivalence theory.

SECOND MAIN THEOREM. *Let L be unimodular and let l be a modular lattice on the same space as L. Then $l \to L$ if and only if $Q(l) \subseteq Q(L) \subseteq Q(l^{\#})$.*

Proof. The necessity is obvious. Therefore suppose that $Q(l) \subseteq Q(L) \subseteq Q(l^{\#})$; thus also $\mathfrak{g}l \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l^{\#}$. If $\mathfrak{s}l_t = \mathfrak{p}^i$, $Q(l^{\#}) = \pi^{-2i}Q(l)$, whence $i \geqq 0$, i.e. $\mathfrak{s}l \subseteq \mathfrak{s}L$. The theorem follows by Theorem 6.6 if $\dim L \geqq 3$, and is also clear for $\dim L = 1$. If $\dim L = 2$, the theorem results from Theorem 6.12.

7. Representation when the dimensions differ.

THIRD MAIN THEOREM. *Let L be a unimodular lattice of dimension $\geqq 3$, and let l be another lattice with $\nu = \dim L - \dim l > 0$. Then necessary and sufficient conditions for $l \to L$ are:*

(i) *if* $\nu = 1$, $l \perp \langle dl \cdot dL \rangle \to L$.

(ii) *if* $\nu = 2$, $l \perp \langle \alpha \rangle \perp \langle \alpha \cdot dl \cdot dL \rangle \to L$ *where α is a non-zero field element such that* $Fl \perp \langle \alpha \rangle \to FL$ *and* $\alpha \in 4(\mathfrak{s}l^{\#})^{-1}$.

(iii) *if* $\nu \geqq 3$, $\mathfrak{s}l \subseteq \mathfrak{s}L$, $\mathfrak{g}l \subseteq \mathfrak{g}L$.

Remark. The representations in (i) and (ii) can be tested by the results of § 6, e.g., by the First Main Theorem.

Proof of Necessity. Put $\mathfrak{a} = (\mathfrak{s}l^{\#})^{-1}$. If $\nu = 1$, there is a unary (almost) $\mathfrak{a}$-modular lattice k such that $l \perp k \to L$ by Theorem 5.1. It is easy to see that $|dl| \leqq |dk|$, and therefore that $dl \cdot dL \in dk \cdot \mathfrak{o}^2$ by discriminants. Thus $l \perp \langle dl \cdot dL \rangle \to l \perp k \to L$.

Next suppose $\nu = 2$. By Theorem 5.1 there is a binary (almost) $\mathfrak{a}$-modular lattice k such that $l \perp k \to L$. It is clear that $FP \cong Fk$ where $P \cong \langle \alpha \rangle \perp \langle \alpha \cdot dl \cdot dL \rangle$. Moreover $\mathfrak{n}P \subseteq 4\mathfrak{a}$ since $dl \cdot dL \in \mathfrak{o}$. If k is modular, it contains a $4\mathfrak{a}$-maximal lattice M by Lemma 5.4, whence $P \to M \to k$ by Lemma 5.3, and so $l \perp P \to l \perp k \to L$ as required. If k is almost modular, then k is $\mathfrak{a}$-maximal and so $l \perp P \to l \perp k \to L$ again.

The necessity is trivial for $\nu \geqq 3$.

Proof of Sufficiency. The cases $\nu \leqq 2$ are trivial, so suppose $\nu \geqq 3$. By splitting hyperbolic planes off L, we can assume that $\nu = 3$ or 4. By field theory, $Fl \perp V \cong FL$ for some space V of dimension ν. Define $\mathfrak{b} = \mathfrak{a} \cap b\mathfrak{o}$ where b is a base generator of $\mathfrak{g}L$. By 3.14, 3.15 or 3.16, there is an (almost) $\mathfrak{b}$-modular lattice k on V with norm group $\mathfrak{b}$. It is clear that $\mathfrak{g}(l \perp k) \subseteq \mathfrak{g}L \subseteq \mathfrak{g}(l \perp k)^{\#}$ and so $l \to l \perp k \to L$ by Theorem 6.6. This finishes the proof of the Third Main Theorem.

We remark that the only case of different dimensions not covered by the Third Main Theorem is $\dim L = 2$, $\dim l = 1$. But this has been done previously in Theorem 6.12 (or better, Lemma 6.13).

The rest of § 7 will be devoted to alternate solutions for the case $\nu = 2$. One of these, namely Theorem 7.2, actually gives a simpler answer than that in the Third Main Theorem, (ii). The other, Theorem 7.3, is ostensibly more complicated, but certainly is at least as simple from a "practical" point of view.

LEMMA 7.1. *If L is a maximal lattice of dimension $\geqq 2$, then $\mathfrak{g}L = \mathfrak{n}L$.*

Proof. It is easy to see that there is a binary (almost) modular sublattice L' of L which splits L, (consider a Jordan form of L); L' is clearly maximal. It suffices to prove that $\mathfrak{g}L' = \mathfrak{n}L'$. Thus we can assume that $\dim L = 2$.

The consideration of a canonical form of L shows that there are elements α and β in F such that $\mathfrak{n}L = \alpha\mathfrak{o} = \beta\pi^{-1}\mathfrak{o}$, and such that α and $\alpha + \beta \in Q(FL)$ if $\mathfrak{d}(-dL) \not\sim \pi$, or α and $\beta \in Q(FL)$ if $\mathfrak{d}(-dL) \sim \pi$. By 91:3 of [4], $Q(L) = Q(FL) \cap \mathfrak{n}L$, and the result follows immediately.

THEOREM 7.2. *Let L be unimodular, and let l be a lattice such that $\dim l = \dim L - 2$, and $\mathfrak{n}L \subseteq \mathfrak{n}l^{\#}$. Then $l \to L$ if and only if $Fl \to FL$, $\mathfrak{s}l \subseteq \mathfrak{s}L$, and $\mathfrak{g}l \subseteq \mathfrak{g}L$.*

Proof. Put $\mathfrak{s} = \mathfrak{s}l^{\#}$ and choose a rational integer i such that $\mathfrak{p}^{i} \subseteq 4\mathfrak{p}\mathfrak{s}^{-1}$. By Lemma 5.2, there is an (almost) $\mathfrak{p}^{i}$-modular lattice k such that $\mathfrak{g}k \subseteq \mathfrak{g}L$, $Fl \perp Fk \cong FL$. If k is almost modular, then $l \to l \perp k \to L$ by Theorem 6.11 (i). Suppose k is modular. By Lemma 5.4, $k^{\#}$ contains a $4\mathfrak{p}^{-i}$-maximal lattice $M^{\#}$, (we put $M = (M^{\#})^{\#}$). Now $\mathfrak{n}M^{\#} \supseteq \mathfrak{s} \supseteq \mathfrak{m} = \mathfrak{n}(l \perp k)^{\#\mathfrak{s}}$. Also $\mathfrak{g}k_{\mathfrak{m}}{}^{\#} \supseteq \mathfrak{g}M_{\mathfrak{m}}{}^{\#} \supseteq \mathfrak{p}\mathfrak{m}$ by Lemma 7.1 since $M^{\#}$ contains an $\mathfrak{m}$-maximal lattice M', and clearly $M_{\mathfrak{m}}{}^{\#} \supseteq M'$. Thus $\mathfrak{g}(l \perp k)_{\mathfrak{m}}{}^{\#} = \mathfrak{g}l^{\#} + \mathfrak{g}k_{\mathfrak{m}}{}^{\#} \supseteq \mathfrak{m} \supseteq \mathfrak{g}L$, and so $l \to l \perp k \to L$ by Theorem 6.7.

THEOREM 7.3. *Let L be unimodular, l a lattice with $\dim l = \dim L - 2$, $\mathfrak{s}l \subseteq \mathfrak{s}L$, and $\mathfrak{n}l^{\#} \subset \mathfrak{n}L$. Suppose that $Fl \perp V \cong FL$ where V is a binary space with $(dV)^{-1}\mathfrak{d}(-dV) = \delta\mathfrak{o}$. Let a be a norm generator of $\mathfrak{g}L$, and define α as*

$$\alpha = \begin{cases} a + 4\delta^{-1}a & \text{if } a \notin Q(V), \\ a & \text{if } a \in Q(V). \end{cases}$$

Then $l \to L$ if and only if $|a| \in |Q(V)|$ and

$$\mathfrak{g}l + \mathfrak{g} \subseteq \mathfrak{g}L \subseteq \mathfrak{g}l^{\#} + \pi^{-2i}\mathfrak{g}$$

where $\mathfrak{s}l^{\#} = \mathfrak{p}^{-i}$ and $\mathfrak{g}$ is the norm group generated by $\pi^{2i}\alpha$ and $\alpha^{-1}\delta$.

Remarks. 1). The conditions of $\mathfrak{s}l \subseteq \mathfrak{o}$ and $\mathfrak{n}l^{\#} \subset \mathfrak{o}$ imply that no component l_i in a Jordan splitting of l has $\mathfrak{n}l_i = \mathfrak{s}l_i$; thus l is the direct sum of binary modular lattices, whence $dl \sim 1$, whence $dV \sim dl \cdot dL \sim 1$. Thus $\delta \in \mathfrak{p}$.

2). The condition $|a| \in |Q(V)|$ is easily seen to be equivalent to the fact that V is not isometric to $\langle a\pi\rangle \perp \langle -\Delta a\pi\rangle$. Thus if this condition holds and $a \notin Q(V)$, then $\delta \notin 4\mathfrak{o}$ and so $|\alpha| = |a|$.

Proof of Necessity. By Theorem 5.1, $l \perp R \to L$ for some binary $\mathfrak{p}^i$-modular lattice R. Thus we also have $L \to (l \perp R)_{\mathfrak{m}}{}^{\#} = l^{\#} \perp R_{\mathfrak{m}}{}^{\#}$ where $\mathfrak{m} = \mathfrak{n}L \supset \mathfrak{n}l^{\#}$. The lattice $R_{\mathfrak{m}}$ is modular by Lemma 4.4, and $\mathfrak{s}R \subseteq \mathfrak{s}R_{\mathfrak{m}} = \mathfrak{p}^j$, say. It is evident that

$$2\mathfrak{s}R_{\mathfrak{m}}{}^{\#} \subseteq 2\mathfrak{s}l^{\#} \subset \mathfrak{n}L \subseteq \mathfrak{n}R_{\mathfrak{m}}{}^{\#};$$

from this it follows that $\mathfrak{n}L = \mathfrak{n}R_{\mathfrak{m}}{}^{\#}$ by Lemma 4.5. Therefore $|a| \in |Q(V)|$, and also we can write $R_{\mathfrak{m}} \cong \mathcal{B}_j(\alpha_0\pi^{2j}, \alpha_0{}^{-1}\delta')$ where $|\alpha_0| = |\alpha| = |a|$, and $dV \cong -1 + \delta'$, $|\delta'| = |\delta|$.

By the choice of α and Lemma 3.7, $\mathfrak{d}(a\alpha) \subseteq \mathfrak{d}(a\alpha_0)$ and therefore $\alpha_0\pi^{2j} \in \mathfrak{g}L$ implies that $a\pi^{2i} \in \mathfrak{g}L$, since $j \leqq i$. If $\alpha_0\delta' \in 2\mathfrak{p}^j$, then $\alpha^{-1}\delta \in 2\mathfrak{p}^j \subseteq \mathfrak{g}L$ since $|\alpha_0{}^{-1}\delta'| = |\alpha^{-1}\delta|$. If $\alpha_0{}^{-1}\delta' \notin 2\mathfrak{p}^j$, then $\alpha_0{}^{-1}\delta' \sim b$ where b is a base generator of $\mathfrak{g}L$; therefore $\alpha^{-1}\delta \in \alpha_0{}^{-1}\delta'\mathfrak{o} \subseteq b\mathfrak{o}$ since $\mathfrak{g}R_{\mathfrak{m}} \subseteq \mathfrak{g}L$. Therefore $\mathfrak{g} \subseteq \mathfrak{g}L$ and so $\mathfrak{g}l + \mathfrak{g} \subseteq \mathfrak{g}L$.

Next let us define $R' \cong \mathcal{B}_{-j}(\alpha, \alpha^{-1}\delta\pi^{-2j})$. Clearly $\mathfrak{g}l^{\#} + \mathfrak{g}R' \subseteq \mathfrak{g}l^{\#} + \pi^{-2i}\mathfrak{g}$. Now $\mathfrak{g}L \subseteq \mathfrak{g}l^{\#} + \mathfrak{g}R_{\mathfrak{m}}{}^{\#}$ and $\mathfrak{g}l^{\#} + \mathfrak{g}R_{\mathfrak{m}}{}^{\#} = \mathfrak{g}l^{\#} + \mathfrak{g}R'$ by Lemma 6.8 (note that a is a norm generator for $\mathfrak{g}(l \perp R_{\mathfrak{m}})^{\#}$ since $a \in \mathfrak{g}L \subseteq \mathfrak{g}(l \perp R_{\mathfrak{m}})^{\#}$ and $\mathfrak{n}(l \perp R_{\mathfrak{m}})^{\#} = \alpha\mathfrak{o} = a\mathfrak{o}$). Therefore $\mathfrak{g}L \subseteq \mathfrak{g}l^{\#} + \pi^{-2i}\mathfrak{g}$.

Proof of Sufficiency. Since $|a| \in |Q(V)|$ and by Lemma 3.7, there is an $\alpha_1 \in Q(V)$ such that $|\alpha_1| = |a|$ and $\mathfrak{d}(a\alpha_1) = \mathfrak{d}(a\alpha)$. Let $v \geqq 0$ be the smallest integer such that $\alpha_1{}^{-1}\delta'\pi^{2v} \in \mathfrak{p}^i$, where $dV \cong -1 + \delta'$ as in the proof of the necessity. If we define $P \cong \mathcal{A}_i(\pi^{2i-2v}\alpha_1, \alpha_1{}^{-1}\delta'\pi^{2v})$, then $F(l \perp P) \cong FL$. We shall show that $l \perp P \to L$.

Suppose first that $\mathfrak{n}P = \alpha_1\mathfrak{p}^{2i}$, so that in particular, $v = 0$. If we define $P' \cong \mathcal{A}_i(\pi^{2i}\alpha, \alpha^{-1}\delta)$, then $\mathfrak{g}(l \perp P')^{\#} = \mathfrak{g}l^{\#} + \pi^{-2i}\mathfrak{g}$, and furthermore $\mathfrak{g}(l \perp P')^{\#} = \mathfrak{g}(l \perp P)^{\#}$ by Lemma 6.8 since a is a norm generator of $\mathfrak{g}l^{\#} + \pi^{-2i}\mathfrak{g}$. Therefore $\mathfrak{g}L \subseteq \mathfrak{g}(l \perp P)^{\#}$ and also

$$\mathfrak{g}P \subseteq \pi^{2i}(\mathfrak{g}l^{\#} + \pi^{-2i}\mathfrak{g}) \subseteq \mathfrak{g}l + \mathfrak{g}$$

since $\pi^i l^{\#} \subseteq l$. Hence $\mathfrak{g}(l \perp P) \subseteq \mathfrak{g}L$.

Now suppose that $\mathfrak{n}P \supset \alpha_1\mathfrak{p}^{2i}$. Then a straightforward calculation shows that $\mathfrak{g}P \supset \alpha_1\mathfrak{p}^{2i}$ whence $\mathfrak{g}L \subseteq \mathfrak{n}L \subseteq \mathfrak{g}P^{\#}$. It is also clear that $\delta \sim \pi$, so that $\alpha^{-1}\delta\mathfrak{o} \subset \mathfrak{g}L$ since $\mathfrak{g} \subseteq \mathfrak{g}L$. But $\mathfrak{n}P \supset \alpha_1\mathfrak{p}^{2i}$ also implies that $\mathfrak{n}P \subseteq \alpha^{-1}\delta\mathfrak{o}$; hence $\mathfrak{g}(l \perp P) \subseteq \mathfrak{g}L$.

We have therefore shown that $F(l \perp P) \cong FL$ and that $\mathfrak{g}(l \perp P) \subseteq \mathfrak{g}L \subseteq \mathfrak{g}(l \perp P)^{\#}$. It follows that $l \perp P \rightarrow L$ by Theorem 6.6, and so the proof of Theorem 7.3 is finished.

THEOREM 7.4. *If L is unimodular, and $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$, then*

$$Q(L) = \begin{cases} dL \cdot \mathfrak{o}^2 & \textit{if } \dim L = 1 \\ \{\alpha \in \mathfrak{g}L \mid (\mathfrak{d}(a\alpha))^2 \subseteq 4\alpha\mathfrak{n}L + (\mathfrak{d}(-dL))^2\} \cap Q(FL) & \textit{if } \dim L = 2 \\ \mathfrak{g}L \cap Q(FL) & \textit{if } \dim L = 3 \\ \mathfrak{g}L & \textit{if } \dim L \geqq 4 \end{cases}$$

Proof. The theorem is obvious if $\dim L = 1$. If $\dim L = 2$, it results from Lemma 6.13. Theorem 7.2 proves it for $\dim L = 3$, and the Third Main Theorem, (iii), takes care of $\dim L \geqq 4$.

8. Concluding remarks. What can be said about the general problem of $l \rightarrow L$ where L is not necessarily modular?

First suppose that $\nu = \dim L - \dim l > 0$. If $\nu = 1$ or 2, then the "reduction" to $\nu = 0$ given for a unimodular lattice L in (i) and (ii) of the Third Main Theorem of § 7 are also valid in this more general setting, if in (ii) the ideal $4(\mathfrak{s}l^{\#})^{-1}$ is replaced by $4(\mathfrak{s}l^{\#})^{-1} \cap 4(\mathfrak{s}L^{\#})^{-1}$, (and if $\mathfrak{s}L \subseteq \mathfrak{o}$). A solution for $\nu \geqq 3$ would depend more directly on the case $\nu = 0$.

Let $L = L_1 \perp \cdots \perp L_T$ be a quasi-Jordan splitting and $l = l_1 \perp \cdots \perp l_t$ be a quasi-Jordan splitting. Define $\mathfrak{L}_j = 0$ if $\mathfrak{s}L \subset \mathfrak{p}^j$, otherwise put $\mathfrak{L}_j = L_1 \perp \cdots \perp L_r$ where r is the largest integer such that $\mathfrak{p}^j \subseteq \mathfrak{s}L_r$. Define $\mathfrak{l}_j$ similarly, using the splitting $l = l_1 \perp \cdots \perp l_t$. Suppose that $\phi : l \rightarrow L$, and that $\mathfrak{s}l_t \subset \mathfrak{p}^j \subseteq \mathfrak{s}l$. Define the linear mapping $\psi : \mathfrak{l}_j \rightarrow \mathfrak{L}_j$ by $\phi(x) = \psi(x) + x'$ where $x \in \mathfrak{l}_j$, $\psi(x) \in \mathfrak{L}_j$, and $x' \in \mathfrak{L}_j^{*}$. It is easy to see that ψ is injective ($\ker \psi = 0$) and so $\dim \mathfrak{l}_j \leqq \dim \mathfrak{L}_j$. Thus a necessary condition for $l \rightarrow L$ is that $\dim \mathfrak{l}_j \leqq \dim \mathfrak{L}_j$ for all values of j. The following theorem is a generalization of Theorem 6.6.

THEOREM. *Let l and L be lattices on the same space (in particular, $\dim l = \dim L$) and suppose that*

$$\begin{aligned} &\dim \mathfrak{L}_j \geqq \dim \mathfrak{l}_j + 3 \ \textit{for} \ \operatorname{ord}(\mathfrak{s}L) \leqq j < \operatorname{ord}(\mathfrak{s}l_t) \\ &\dim \mathfrak{L}_j \geqq \dim \mathfrak{l}_j \ \textit{otherwise.} \end{aligned} \tag{6}$$

Then $l \rightarrow L$ if and only if $\mathfrak{g}l \subseteq \mathfrak{g}L$ and $\mathfrak{g}L_{\mathfrak{s}}^{\#} \subseteq \mathfrak{g}l_{\mathfrak{s}}^{\#}$, where $\mathfrak{s} = (\mathfrak{s}l_t)^{-1}$.

Proof. The proof of the necessity is straightforward and is omitted. The proof of the sufficiency proceeds by induction on $N(l, L) = T \cdot \operatorname{ord}(\mathfrak{s}l_t)$. By scaling l and L if necessary, we may assume that $\mathfrak{s}L = \mathfrak{o}$. Putting

$j = \operatorname{ord}(\mathfrak{s}l)$ in (6), we see that $\mathfrak{s}l \subseteq \mathfrak{s}L$, so that $N(l, L) \geqq 0$. Also $\dim L \geqq \dim L_1 \geqq 3$ follows from (6) for $j = \operatorname{ord}(\mathfrak{s}L)$. If $N(l, L) = 0$, then l and L are both unimodular or almost unimodular and so $l \cong L$ since $\mathfrak{g}l = \mathfrak{g}L$. Suppose then that $N(l, L) > 0$. Let $\mathfrak{s}l_t = \mathfrak{p}^i$. Condition (6) for $j = i$ shows that $\mathfrak{s}l_t \subseteq \mathfrak{s}L_t$ since $\mathfrak{l}_t = l$ and $\dim l = \dim L$.

First Case: $\mathfrak{s}l_t \subset \mathfrak{s}L_t$. The induction step for this case is almost the same as that in Theorem 6.6. By the same theorem, we can suppose that $T > 1$, whence $\mathfrak{s}l_t \subseteq \mathfrak{p}^2$. Now $\dim l_t \geqq 3$ by (6) for $j = \operatorname{ord}(\mathfrak{s}l_t) - 1$, Thus we can suppose that the quasi-Jordan splitting for l has been chosen so that $\mathfrak{g}(\pi^{-i}l_t) = \mathfrak{g}l_{\mathfrak{s}}{}^{\#}$. Then we find an (almost) $\mathfrak{p}^{i-1}$-modular lattice k on Fl_t with $\mathfrak{g}k = \mathfrak{g}L \cap \mathfrak{p}^{i-1} \cap \pi^{-2}\mathfrak{g}l_t$. Put $J = l_t{}^* \perp k$. Then $l \rightarrow J$ since $l_t \rightarrow k$, and $\mathfrak{g}J \subseteq \mathfrak{g}L$. If $\mathfrak{s}' = \mathfrak{p}^{-i+1} = (\mathfrak{s}k)^{-1}$, then

$$\begin{aligned} \mathfrak{g}J_{\mathfrak{s}'}{}^{\#} \supseteq \mathfrak{g}(\pi^{-i+1}k) &= \pi^{-2(i-1)}\mathfrak{g}k = (\pi^{-2(i-1)}\mathfrak{g}L) \cap \mathfrak{p}^{-i+1} \cap \pi^{-2i}\mathfrak{g}l_t \\ &\supseteq \mathfrak{g}L_{\mathfrak{s}'}{}^{\#} \cap \mathfrak{g}l_{\mathfrak{s}}{}^{\#} = \mathfrak{g}L_{\mathfrak{s}'}{}^{\#}. \end{aligned}$$

Thus $\mathfrak{g}L_{\mathfrak{s}'}{}^{\#} \subseteq \mathfrak{g}J_{\mathfrak{s}'}{}^{\#}$. But J inherits the properties analogous to (6) from l, and $N(J, L) < N(l, L)$. Therefore $J \rightarrow L$ by induction, and so $l \rightarrow J \rightarrow L$ as required.

Second Case: $\mathfrak{s}l_t = \mathfrak{s}L_T$. We shall assume again that $\mathfrak{g}(\pi^{-i}l_t) = \mathfrak{g}l_{\mathfrak{s}}{}^{\#}$; since $\dim L_1 \geqq 3$, a similar argument allows us to assume that $\mathfrak{g}L_1 = \mathfrak{g}L$. Condition (6) shows that $\dim l_t \geqq \dim L_T + 3$. The reader can easily check that (iii) of the Third Main Theorem remains valid if L is almost unimodular instead of unimodular. Therefore $L_T \rightarrow l_t$ since

$$\mathfrak{g}L_T = \pi^{2i}\mathfrak{g}L_{T\mathfrak{s}}{}^{\#} \subseteq \pi^{2i}\mathfrak{g}l_{\mathfrak{s}}{}^{\#} = \pi^{2i}\mathfrak{g}(\pi^{-i}l_t) = \mathfrak{g}l_t,$$

whence $l_t \cong L_T \perp J$ for some (almost) modular lattice J of dimension $\geqq 3$. But FJ supports an (almost) $\mathfrak{p}^i$-modular lattice K with norm group $\mathfrak{g}K = \mathfrak{g}J + \mathfrak{g}l_t = \mathfrak{g}l_t$; this can be proved by first applying one of 3.12 and the necessity of 3.15 and 3.16, followed by one of 3.14 and the sufficiency of 3.15 and 3.16. By equality of norm groups, $l_t \cong L_T \perp K$. Define $l' = l_t{}^* \perp K$ and $L' = L_T{}^*$. Clearly l' and L' satisfy the conditions analogous to (6), and $Fl' \cong FL'$ by Witt's Theorem. Also

$$\begin{gathered} \mathfrak{g}l' = \mathfrak{g}l \subseteq \mathfrak{g}L = \mathfrak{g}L_1 = \mathfrak{g}L' \\ \mathfrak{g}L'^{\#} \subseteq \mathfrak{g}L^{\#} \subseteq \mathfrak{g}l_{\mathfrak{s}}{}^{\#} = \mathfrak{g}(\pi^{-i}l_t) = \mathfrak{g}(\pi^{-i}K) = \mathfrak{g}l_{\mathfrak{s}}'^{\#}. \end{gathered}$$

Since $N(l', L') < N(l, L)$, $l' \rightarrow L'$ by induction. Therefore $l \cong l' \perp L_t \rightarrow L' \perp L_t = L$ as required. This finishes the proof of the theorem.

COROLLARY. *Let l and L be two lattices such that*

$$\dim \mathfrak{L}_j \geqq \dim \mathfrak{l}_j + 3 \text{ for } j \geqq \operatorname{ord}(\mathfrak{s}L), \tag{7}$$

(*in particular,* $\dim L - \dim l \geqq 3$). *Then* $l \to L$ *if and only if* $\mathfrak{s}l \subseteq \mathfrak{s}L$ *and* $\mathfrak{g}l \subseteq \mathfrak{g}L$.

Proof. The proof is omitted. It is entirely similar to the proof of the Third Main Theorem, (iii), with the above theorem being used instead of Theorem 6.6.

In this paper, we have given a complete solution to the problem $l \to L$ when L is modular, and only a partial solution when L is not modular. The theory remaining to be developed is as follows, where $\nu = \dim L - \dim l$, and L is not modular:

$\nu = 0$: L does not satisfy (6)
$\nu \geqq 3$: L does not satisfy (7).

Of course, the "reduction" to $\nu = 0$ given earlier for $\nu = 1$ or 2 will be a solution only when the case $\nu = 0$ is completely solved.

McGILL UNIVERSITY,
UNIVERSITY OF NOTRE DAME.

REFERENCES.

[1] H. Hasse, "Symmetrische Matrizen im Körper der rationalen Zahlen," *Journal für die Reine und Angewandte Mathematik*, vol. 153 (1924), pp. 12-43.

[2] B. W. Jones, *The Arithmetic Theory of Quadratic Forms*, Carus Monographs, No. 10, Mathematical Association of America, 1950.

[3] O. T. O'Meara, "The integral representations of quadratic forms over local fields," *American Journal of Mathematics*, vol. 80 (1958), pp. 843-878.

[4] ———, *Introduction to Quadratic Forms*, Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 1963.

QUASICONFORMAL MAPPINGS AND EXTREMAL LENGTHS.

By H. RENGGLI.*

1. Introduction. In our attempt to study homeomorphisms which preserve curve families with extremal length zero, we found the rather surprising result that these mappings are quasiconformal in the geometrical sense (Definition 1′ and Theorem 1′).

In selecting special families of curves with extremal length zero, we have been able to give a full characterization of quasiconformal mappings by homeomorphisms which leave exactly these selected families invariant (Definition 1 and Theorem 1).

Hence this property can be used as a definition of quasiconformal mappings. Furthermore it can advantageously be applied to the study of invariance and continuation properties of quasiconformal mappings.

We would like to draw attention to the following fact only: A. Beurling and L. Ahlfors [2] have pointed out that under a quasiconformal mapping of a half plane onto itself—which implies a homeomorphism of the boundary line—neither the class of sets of Lebesgue measure zero nor the class of sets of harmonic measure zero situated on the line is preserved.

Therefore it is of some interest to give an example of a class of null-sets which is preserved by the induced homeomorphism of the boundary (Theorem 2).

2. Two lemmas for extremal lengths. The definition and the main properties of extremal lengths have been given by L. Ahlfors and A. Beurling [1]. An account of the theory of extremal lengths can e. g. be found in [7].

The first lemma will be proved for families of curves in the complex plane; but it can readily be generalized to families on a given Riemann surface.

LEMMA 1. *For each* i $(i=1,2,3,\cdots)$, Γ_i *is a family of curves* γ_i *lying in* G_i, *where the* G_i *are disjoint measurable sets. Let* $\Gamma = \bigcup \Gamma_i$. *Then*

$$\sum_{i=1}^{\infty} \frac{1}{\lambda(\Gamma_i)} = \frac{1}{\lambda(\Gamma)}.$$

Received August 3, 1962.

* Research supported by the National Science Foundation, Grant NSF G-21219.

Proof. (I) Since $\Gamma_i \subset \Gamma$, we have $\lambda(\Gamma_i) \geqq \lambda(\Gamma)$ for every i. If $\lambda(\Gamma_i) = 0$ for some i, then the lemma trivially holds. Let therefore $\lambda(\Gamma_i) \neq 0$ for every i. To every given $\epsilon_i > 0$ there is a ρ_i defined on G_i such that $L_{\Gamma_i}(\rho_i) \geqq 1$ and $F_{\Gamma_i}(\rho_i) \leqq \dfrac{1}{\lambda(\Gamma_i)} + \epsilon_i$. But ρ defined by $\rho = \rho_i$ in G_i and zero elsewhere is admissible for Γ and $L_\Gamma(\rho) \geqq 1$. Furthermore

$$\frac{1}{\lambda(\Gamma)} \leqq F_\Gamma(\rho) = \sum_{i=1}^{\infty} F_{\Gamma_i}(\rho_i) \leqq \sum_{i=1}^{\infty} \frac{1}{\lambda(\Gamma_i)} + \epsilon,$$

where $\epsilon = \sum_{i=1}^{\infty} \epsilon_i$. Hence $\dfrac{1}{\lambda(\Gamma)} \leqq \sum_{i=1}^{\infty} \dfrac{1}{\lambda(\Gamma_i)}$, and the lemma trivially holds if $\lambda(\Gamma) = 0$.

(II) Let $\lambda(\Gamma) \neq 0$. To every $\eta > 0$ there is a ρ such that $L_\Gamma(\rho) \geqq 1$ and $F_\Gamma(\rho) \leqq \dfrac{1}{\lambda(\Gamma)} + \eta$. Let ρ_i be the restriction of ρ to G_i. Then $\sum_{i=1}^{\infty} \dfrac{1}{\lambda(\Gamma_i)} \leqq \sum_{i=1}^{\infty} F_{\Gamma_i}(\rho_i) \leqq F_\Gamma(\rho)$ and hence $\sum_{i=1}^{\infty} \dfrac{1}{\lambda(\Gamma_i)} \leqq \dfrac{1}{\lambda(\Gamma)}$.

The second lemma deals with families in the complex z-plane.

LEMMA 2. *Let M be a family of curves such that each curve μ given by $z_\mu(t)$, continuous on $0 \leqq t \leqq 1$, satisfies $\Im z_\mu(t) > 0$ for $0 < t < 1$ and $\Im z_\mu(0) = \Im z_\mu(1) = 0$.*

Let $N = \{\nu\}$ be the reflection of $\{\mu\}$ in the real axis and let Γ be the family of all closed curves γ such that γ contains exactly one curve μ and exactly one curve ν.

Then $\lambda(M) = \lambda(N) = \tfrac{1}{2}\lambda(\Gamma)$.

Proof. (I) That $\lambda(M) = \lambda(N)$ is obvious by symmetry. From the definition of Γ one gets, by a known property of extremal lengths, $\lambda(M) + \lambda(N) \leqq \lambda(\Gamma)$.

(II) Let $\lambda(\Gamma) \neq \infty$. Then to every $\epsilon > 0$ there is a ρ such that $F_\Gamma(\rho) = 1$ and $\lambda(\Gamma) - \epsilon \leqq L^2_\Gamma(\rho)$.

Let $\rho = \rho(z)$, $\tilde\rho = \rho(\bar z)$ and $\rho = \tfrac{1}{2}[\rho + \tilde\rho]$ for every z. By an application of Schwarz's inequality one gets $F_\Gamma(\tilde\rho) \leqq 1$. Since $\tilde\rho$ is symmetric to the real axis, $L_\Gamma(\tilde\rho)$ can be approximated by choosing curves also symmetric with respect to the real axis. Hence one concludes easily that $L_\Gamma(\tilde\rho) \geqq L_\Gamma(\rho)$.

Let $\hat\rho$ be the restriction of $\tilde\rho$ to the upper half plane. $\hat\rho$ is admissible for M and $F_M(\hat\rho) = \tfrac{1}{2}F_\Gamma(\tilde\rho)$, $L_M(\hat\rho) = \tfrac{1}{2}L_\Gamma(\tilde\rho)$. Therefore

$$\lambda(\Gamma) - \epsilon \leqq L^2_\Gamma(\tilde\rho) \leqq 2\,\frac{L^2_M(\hat\rho)}{F_M(\hat\rho)} \leqq 2\lambda(M).$$

Similarly one shows that if $\lambda(\Gamma)=\infty$, then $\lambda(M)=\infty$.

3. A characterization of quasiconformal mappings.

A quadrilateral Q (in the plane) is the interior of a Jordan curve C on which four distinct points have been selected. These four points divide C into four arcs, the sides of Q. A quadrilateral on a Riemann surface is defined by a topological image of $Q \cup C$ or simply of Q.

Definition 1. Let Q_i be a quadrilateral on a Riemann surface. The family Γ_i of all curves γ_i, $\gamma_i \subset Q_i$, which join two opposite sides of Q_i, is called a simple quadrilateral family. Any (necessarily countable) union $\Gamma = \cup \Gamma_i$ of simple quadrilateral families such that the corresponding Q_i are mutually disjoint is called a quadrilateral family of curves.

THEOREM 1. *A topological mapping f of one Riemann surface onto another is quasiconformal in the geometrical sense if and only if the set of quadrilateral families whose extremal length is zero is invariant under f; i. e. for any quadrilateral family Γ the extremal length $\lambda(\Gamma)=0$ if and only if $\lambda(\Gamma')=0$ where Γ' denotes the image of Γ under f.*

Proof. For each simple quadrilateral family Γ_i let m_i be the modulus reciprocal to $\lambda(\Gamma_i)$. Applying Lemma 1 one immediately concludes that Theorem 1 is equivalent to the main theorem in [8], whose generalization to mappings of Riemann surfaces is quite trivial.

Definition 1′. We denote by f^* a homeomorphism of one Riemann surface onto another such that the set of families of curves whose extremal length vanishes is invariant under f^*; i. e. for any family Γ of curves, $\lambda(\Gamma)=0$ if and only if $\lambda(\Gamma')=0$ where Γ' denotes the image of Γ under f^*.

THEOREM 1′. *Every differentiable quasiconformal mapping is an f^* and every f^* is quasiconformal in the geometrical sense.*

Proof. Let Γ be a family of curves and let Γ' denote its image under a given differentiable quasiconformal mapping. Then it is well known that there is a constant $K \geqq 1$ such that $\frac{1}{K}\lambda(\Gamma) \leqq \lambda(\Gamma') \leqq K\lambda(\Gamma)$ for any Γ. Hence the first statement of Theorem 1′ holds, whereas the second is an immediate consequence of Theorem 1.

4. Some invariance and continuation properties.

Most of the results in this section have been announced in *Bulletin of the American Mathematical Society*, vol. 63 (1957), pp. 273-274 and *Notices of the American Mathe-*

matical Society, vol. 5 (1958), pp. 148-149. But also other authors [5], [6], [9], [10] have treated the problems independently. We nevertheless give a summary hoping to have unified and simplified some of the questions.

Let G be a domain in the Riemann sphere R. Any boundary component of G is either a point denoted by σ_G or a continuum. By a continuum we mean a compact connected set consisting of more than one point. We will only deal with the set $B_G = \{\sigma_G\}$.

H. Grötzsch [4] was the first to start to classify B_G by giving

Definition 2. σ_G is called "completely point-like" (with respect to G) and denoted by $\sigma_G{}^*$ if and only if for every conformal mapping of G into R, σ_G corresponds (in the induced boundary correspondence) to a point (and hence never to a continuum).

Using conformal mapping onto a minimal circular slit domain one proves (preferably by contraposition)

PROPOSITION 1. *Let C be a Jordan curve, $C \subset G$. C decomposes R into three parts, R', R'' and C. Let $\sigma_G \in R'$ and $G' = R' \cap G$. Let Γ be the family of all closed curves γ, $\gamma \subset G'$, such that each γ separates σ_G from C (with respect to R). Then σ_G is a $\sigma_G{}^*$ if and only if $\lambda(\Gamma) = 0$* [5].

It can be shown that $\lambda(\Gamma) = 0$ is independent of the special choice of C, and that there is a Γ^*, $\Gamma^* \subset \Gamma$, $\lambda(\Gamma^*) = 0$, which can be considered as a quadrilateral family (according to Definition 1). Hence we can apply Theorem 1. Using these remarks one gets successively:

COROLLARY 1. *Let U be a ϵ-neighborhood (in the chordal distance on R) of a given σ_G and let V be that component of $G \cap U$ whose boundary contains σ_G. Then σ_G is a $\sigma_G{}^*$ if and only if σ_V $(= \sigma_G)$ is a $\sigma_V{}^*$; i. e. being "completely point-like" is a local property.*

COROLLARY 2. *Let $G \cup B_G$ be a domain. Using the terminology of* [1], *every σ_G is a $\sigma_G{}^*$ if and only if every compact subset of B_G is a set N_{SB}* [10].

COROLLARY 3. *If σ_G is a $\sigma_G{}^*$, then it is also "completely point-like" with respect to the larger class of quasiconformal mappings. Especially the sets N_{SB} are quasiconformally invariant.*

COROLLARY 4. *Any quasiconformal mapping of G can be continued to a topological mapping of $G \cup B_G$ if and only if any σ_G is a $\sigma_G{}^*$. If the continuation exists, then it is unique* [6]. *Hence this property could have been used to define a $\sigma_G{}^*$.*

Another step in classifying B_G is done by

Definition 3. Let $G \cup B_G$ be a domain. B_G is called "conformally [resp. quasiconformally] removable" (with respect to G) if and only if every conformal [quasiconformal] mapping of G into R can be extended to a conformal [quasiconformal] mapping of $G \cup B_G$.

This is just another way of defining the sets N_D, since by using Corollary 4 and [1], Theorem 5 and § 5, one can readily show that any compact subset of B_G is a set N_D if and only if it is "conformally removable".

COROLLARY 5. *Let $U, U \subset G \cup B_G$ be a ϵ-neighborhood (in the chordal distance of R) of a given σ_G. Then B_G is "conformally removable" if and only if for every σ_G the set $B_V = B_G \cap U$ is "conformally removable" (with respect to $V = G \cap U$); i.e. being "conformally removable" is a local property.*

The following result is a sharpening of K. Strebel's Theorem 1 [12].

PROPOSITION 2. *A set B_G is "quasiconformally removable" if and only if it is "conformally removable".*

Proof. (I) That conformal removability implies quasiconformal removability follows from [12].

(II) Let N be a compact subset of B_G and let f map $R - N$ conformally onto a minimal parallel slit domain. By hypothesis the restriction of f to G can be extended to a quasiconformal mapping $\bar{f}$ of $G \cup B_G$. We note that $\bar{f}(N)$ has area zero. Hence also N has area zero, since the quasiconformal $\bar{f}^{-1}$ is measurable by [3]. Using the hypothesis, the measurability property and [1], Theorem 4, one concludes that N is a set N_D and therefore by the remark following Definition 3 that B_G is "conformally removable".

From Proposition 2 follows

COROLLARY 6. *The sets N_D are quasiconformally invariant* [9].

As a final remark we want to mention that Definition 2 does not imply Definition 3 ([1], Theorem 16); i.e. topological continuability does not imply conformal continuability.

5. Boundary correspondence. That a quasiconformal mapping of a Jordan domain onto another can be extended to a homeomorphism of the closure is a classical result of quasiconformal mappings. Nevertheless we give an outline of another proof in order to illustrate the applicability of our characterization of quasiconformal mappings. Furthermore, E. C. Schlesinger [11] has made a study of prime ends using families of curves with extremal length zero. It would be easy to show that his whole study and hence the classical results on prime ends can be carried over to quasiconformal mappings.

PROPOSITION 3. *Let D be a domain (in the Riemann sphere R), bounded by a Jordan curve C, and let f map D quasiconformally onto another domain D', $D' \subset R$, bounded by a Jordan curve C'.*

Then f can be extended to a topological mapping of $D \cup C$ onto $D' \cup C'$.

Proof. Let $K, K \subset D$, be a continuum and $p, p \in C$, be a point. Let Γ be the family of curves which separate p from K with respect to D. Since p is an accessible point, it is easy to show that $\lambda(\Gamma) = 0$, using known properties of extremal lengths. Furthermore Γ will contain a quadrilateral family Γ^*, $\Gamma^* \subset \Gamma$, with $\lambda(\Gamma^*) = 0$. Using Theorem 1, one concludes that $\lambda(\Gamma') = 0$, where Γ' denotes the image of Γ under f. Hence p cannot correspond to any continuum on the Jordan curve C' because otherwise one would get $\lambda(\Gamma') > 0$. The proof of Proposition 3 can now be completed along classical lines.

Finally we prove

THEOREM 2. *Let D be a domain (in the Riemann sphere R), bounded by a Jordan curve C. Let f map D quasiconformally onto the upper half of the complex plane.*

Let N, $N \subset C$, N compact, be such that every point of N is completely point-like (according to Definition 2) with respect to $R-N$, i.e. N is a set N_{SB}.

Then the image N' of N (by Proposition 3) is also "completely point-like" with respect to $R-N'$, i.e. N' is also a set N_{SB}.

Proof. By hypothesis every point p, $p \in N$, is accessible from D. Let K, $K \subset D$, be a continuum, let $q \in K$, and let C^* be an arc, represented by $z(t)$, continuous (in the chordal distance) on $0 \leqq t \leqq 1$ with $z(0) = q$, $z(t) \in D$ for $0 \leqq t < 1$ and $z(1) = p$. According to Proposition 1, the family Γ of all curves γ, separating (with respect to $R-N$) p from an appropriately chosen Jordan curve passing through q satisfies $\lambda(\Gamma) = 0$. Every γ intersects D along certain arcs. Let for any $\gamma \in \Gamma$, ψ be one of these arcs, which also intersect C^*. Then every such ψ will separate q from p with respect to D. By known properties of extremal lengths, $\lambda(\Psi) = 0$ where $\Psi = \{\psi\}$. Either Ψ is already a simple quadrilateral family, or contains a quadrilateral family whose extremal length is zero. Let Ψ denote this family. Using Theorem 1,

$\lambda(\Psi') = 0$, where Ψ' denotes the image of Ψ under f. Applying Lemma 2, Ψ' can be identified with a family M and hence the family of all the curves which separate with respect to $R - N'$ the image of p' of p from a certain Jordan curve on R, has extremal length zero. Hence by Proposition 1, p' is "completely point-like" with respect to $R - N'$.

It is trivial that Theorem 2 can be generalized to mappings f of D onto domains bounded by piecewise analytic arcs.

For quasiconformal mappings of the upper half plane onto itself the question of boundary invariance of all the known function-theoretic null-sets is now settled by means of [2], Theorems 12 and 13 of [1] and our Theorem 2.

ALBUQUERQUE, NEW MEXICO.

REFERENCES.

[1] L. Ahlfors and A. Beurling, "Conformal invariants and function-theoretic null-sets," *Acta Mathematica*, vol. 83 (1950), pp. 101-129.

[2] ———, "The boundary correspondence under quasiconformal mappings," *Acta Mathematica*, vol. 96 (1956), pp. 125-142.

[3] L. Bers, "On a theorem of Mori and the definition of quasiconformality," *Transactions of the American Mathematical Society*, vol. 84 (1957), pp. 78-84.

[4] H. Grötzsch, "Eine Bemerkung zum Koebeschen Kreisnormierungsprinzip," *Berichte über die Verhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig, Mathematisch-Naturwissenschaftliche Klasse*, vol. 87 (1936), pp. 319-324.

[5] M. Jurchescu, "Modulus of a boundary component," *Pacific Journal of Mathematics*, vol. 8 (1958), pp. 791-809.

[6] ———, "Bordered Riemann surfaces," *Mathematische Annalen*, vol. 143 (1961), pp. 264-292.

[7] H. Renggli, "Zur konformen Abbildung auf Normalgebiete," *Commentarii Mathematici Helvetici*, vol. 31 (1956), pp. 5-40.

[8] ———, "Zur Definition der quasikonformen Abbildungen," *Commentarii Mathematici Helvetici*, vol. 34 (1960), pp. 222-226.

[9] H. L. Royden, "On a class of null-bounded Riemann surfaces," *Commentarii Mathematici Helvetici*, vol. 34 (1960), pp. 52-66.

[10] N. Savage, "Weak boundary components of an open Riemann surface," *Duke Mathematical Journal*, vol. 24 (1957), pp. 79-95.

[11] E. C. Schlesinger, "Conformal invariants and prime ends," *American Journal of Mathematics*, vol. 80 (1958), pp. 83-102.

[12] K. Strebel, "On the maximal dilation of quasiconformal mappings," *Proceedings of the American Mathematical Society*, vol. 6 (1955), pp. 903-909.

GENERALIZED AUTOMORPHIC FORMS AND CERTAIN HOLOMORPHIC VECTOR BUNDLES.

By Mikio Ise.

Introduction. Let X be a symmetric bounded domain of N-complex dimensions and let Γ be a properly discontinuous group of holomorphic automorphisms of X with the following properties:

(a) The quotient space $\Gamma\backslash X$ $(=Y)$ is compact.

(b) Any element $\gamma \in \Gamma$ different from the identity has no fixed point in X.

These two conditions imply that Y is a projective algebraic manifold (without singularities) by a theorem of Kodaira [15]. A mapping R of $\Gamma \times X$ into the complex general linear group $GL(m, \boldsymbol{C})$ $(m \geqq 1)$ is called an *automorphic factor,* if R satisfies the relation:

$$R(\gamma\gamma', x) = R(\gamma, \gamma' x) R(\gamma', x)$$

for γ, $\gamma' \in \Gamma$ and $x \in \gamma$, and further if $R(\gamma, x)$ is holomorphic with respect to the variable $x \in X$ for any fixed $\gamma \in \Gamma$. By an *automorphic form* f with respect to Γ for the automorphic factor R, we mean the holomorphic mapping f of X into the m-dimensional complex cartesian space $\boldsymbol{C}^m$ which satisfies the equalities:

$$f(\gamma, x) = R(\gamma, x) f(x),$$

for $\gamma \in \Gamma$ and $x \in X$.

On the other hand, to such an automorphic factor R corresponds an (m-dimensional) holomorphic vector bundle E_R over $Y = \Gamma\backslash X$ in the following fashion: In fact, Γ acting on the product manifold $X \times \boldsymbol{C}^m$ such that

$$\gamma \circ (x, \xi) = (\gamma \cdot x, R(\gamma, x)\xi)$$

for $\gamma \in \Gamma$ and $(x, \xi) \in X \times \boldsymbol{C}^m$, we obtain the vector bundle E_R as the quotient complex manifold of $X \times \boldsymbol{C}^m$ by Γ. Then, it is readily seen that an automorphic form f with respect to Γ for automorphic factor R can be regarded

---

Received April 16, 1963.

as a holomorphic section of the vector bundle E_B, and conversely. Therefore the space of these automorphic forms will be denoted by $H^0(Y, \boldsymbol{E}_B)$.[1]

The main purpose of the present paper is first to generalize the concept of the automorphic forms of type ρ introduced in [7] in the case of Siegel's symplectic space to the general symmetric bounded domains, and then to compute the dimension of the space of these automorphic forms under the conditions (a) and (b). We shall now explain the contents of the paper:

Let G^c be a connected complex semi-simple Lie group and G^u a compact form (a maximal compact subgroup) of G^c. Further we denote by U a connected, closed complex subgroup of G^c containing no closed normal non-discrete subgroups of G^c; for which we assume that the complex manifold $X^u = G^c/U$ is the compact symmetric hermitian manifold dual to X in the sense of E. Cartan. Then G^u/K $(K = U \cap G^u)$ is a symmetric homogeneous space for a suitable G^u and is naturally identified with X^u; moreover we know that there is a connected, non-compact real semi-simple subgroup G of G^c containing K as a maximal compact subgroup of G. The (symmetric) homogeneous space G/K thus obtained may be identified with X.[2] In the sequel, K^c will denote the complexification of K in G^c, and ρ a holomorphic (completely reducible) representation of K^c into $GL(m, \boldsymbol{C})$. First, in § 2, we will introduce the mapping $J\colon G \times X \to K^c$ which will be called the *canonical factor of automorphy* (cf. Definition 1),[3] and define the automorphic factor $J_\rho = \rho \circ J$ as the composition of ρ and J. In case X is the Siegel's symplectic space, J is of a well-known form (cf. § 9), and an automorphic form for the factor J_ρ is no other than an automorphic form of type ρ in the sense of [7]. In this section, we shall clarify the significance of J by showing the relation between the bundle E_{J_ρ} (we write also as $E_{\rho\circ J}$) and the homogeneous vector bundle E_{ρ^u}[4] over X^u in the sense of Bott [5] (cf. Proposition 1). Next we introduce the cohomology groups $H^q(Y, \boldsymbol{E}_{J_\rho})$ over Y (resp. $H^q(X^u, \boldsymbol{E}_{\rho^u})$ over X^u) $(0 \leqq q \leqq N)$ with the coefficients in the sheaf of germs of holomorphic sections of E_{J_ρ} (resp. of E_{ρ^u}). The zero-dimensional cohomology group $H^0(Y, \boldsymbol{E}_{J_\rho})$ is the very space of automorphic forms with which we shall be concerned throughout this paper. We consider further the Euler-characteristic $\chi(Y, \boldsymbol{E}_{J_\rho})$ (resp. $\chi(X^u, \boldsymbol{E}_{\rho^u})$) with respect to these cohomology groups. One of the main results in this paper is Theorem 1 (in § **6**) which states the formula:

---

[1] See the Notations at the end of this introduction.

[2] See, for detail, §1.

[3] We have learned this concept from I. Satake, which is originally due to A. Orihara; we are very thankful to them.

[4] For the definition of the homogeneous vector bundle E_{ρ^u}, we refer to §2, 6, and originally to [5].

$$\text{(A)} \qquad \chi(Y, \boldsymbol{E}_{J_\rho}) = \chi(Y)\chi(X^u, \boldsymbol{E}_{\rho^u}),$$

where $\chi(Y)$ denotes thet arithmetic genus of the algebraic manifold Y. This formula is a generalization of Hirzebruch's result (see [10, Satz 1 and Satz 4]), and the proof shall be carried out by generalizing the proportionality in the sense of [10] to the *Chern numbers* of the bundles E_{J_ρ} and those of E_{ρ^u} (cf. Lemma 6 in § **6**).

Our next task is to prove vanishing theorems for the cohomology groups $H^q(Y, \boldsymbol{E}_{J_\rho})$ $(q \geqq 0)$. In fact in § **7** we shall show that, *under a certain condition on* ρ,

$$\text{(B)} \qquad H^q(Y, \boldsymbol{E}_{J_\rho}) = 0 \quad \text{for all } q > 0;$$

This is another main result of this paper (cf. Theorem 2). Combining (A) with (B) we see that

$$\text{(C)} \qquad \dim_C H^0(Y, \boldsymbol{E}_{J_\rho}) = \chi(Y)\chi(X^u, \boldsymbol{E}_{\rho^u}),$$

under the same condition as in (B). Here recall that $\chi(Y)$ is computable, as Hirzebruch [11] showed, if we know the total volume of Y with respect to G-invariant metric in X (cf. § **6**), while that $\chi(X^u, \boldsymbol{E}_{\rho^u})$ can be computed within the representation theory of G^c, according to Bott's results [5] (see for detail, Cor. of Theorem 2). We show in § **7**, moreover, theorems which give some sufficient conditions for $H^0(Y, \boldsymbol{E}_{J_\rho}) = 0$.

To prove these vanishing theorems for the vector bundle E_{J_ρ}, we introduce in § **3** another algebraic manifold $Y_T = \Gamma\backslash G/T$ (T is a maximal torus of K) and the line bundle F_σ over Y_T which is determined by ρ. The computations of $H^q(Y, \boldsymbol{E}_{J_\rho})$ are then reduced to those of the cohomology groups $H^q(Y_T, \boldsymbol{F}_\sigma)$ over Y_T with coefficients in the sheaf of germs of holomorphic sections of F_σ. Then, the vanishing theorems of Kodaira [14] can be applied to $H^q(Y_T, \boldsymbol{F}_\sigma)$, if the explicit form of the Chern class of F_σ are known. (These methods are clearly analogous to Bott's theory of homogeneous vector bundles; see [5]). Thus, our main results (A) and (B) shall be obtained by clarifying the explicit forms of the Chern classes of the vector bundles appeared in the above; § **4** will be devoted to these discussions.

In concluding this introduction, we should like to express our thanks to Prof. Y. Matsushima for frequent helpful conversations with him on the present subject. We are also thankful to Prof. I. Satake for his valuable communication.

Note. After completing the present work, the author was informed that R. P. Langlands has calculated the dimension of the space of automorphic forms of type ρ under the condition (a) without (b) by using Selberg's trace formula and some results due to Harish-Chandra (but, his condition on the representation ρ is somewhat stronger than that of ours in Cor. of Theorem 2, § 7). However, our vanishing theorems are new as far as we know, and these theorems remain valid even if the condition (b) is omitted, as the reader could readily see by making use of Baily's work (*Proceedings of the National Academy of Sciences, U. S. A.*, vol. 40, pp. 804-808) in place of Kodaira's [14].

Notations. i) $\boldsymbol{Z}$, $\boldsymbol{R}$ and $\boldsymbol{C}$ denote the ring of integers, the field of real numbers and that of complex numbers respectively. ii) To denote a fibre bundle, we write as $E = E(X, F, G, \pi)$; where E, X, F, G and π will be the bundle space, the base space, the standard fibre, the structure group and the projection respectively. In the principal case, we omit the fibre without any change in others. iii) For a holomorphic vector bundle E, we denote by $\boldsymbol{E}$ the (analytic) sheaf of germs of local holomorphic sections of E. iv) Lie groups are denoted by the great roman letters G, H, K etc., and their Lie algebras by the corresponding small German letters: $\mathfrak{g}$, $\mathfrak{h}$, $\mathfrak{k}$ etc. v) A and B being two complex Lie groups, $\mathrm{Hom}(A, B)$ is by definition the set of all *holomorphic* homomorphisms of A into B; while, in case A and B are real Lie groups, by $\mathrm{Hom}(A, B)$ is meant the set of all *differentiable* homomorphisms of A into B. In each case, for $\rho \in \mathrm{Hom}(A, B)$, we denote by $\rho \in \mathrm{Hom}(\mathfrak{a}, \mathfrak{b})$ the differential of ρ. iv) $\dotplus$ means the direct sum of vector spaces; for the real subspace $\mathfrak{a}$ of a complex vector space, $\mathfrak{a}^c$ denotes the complex subspace spanned by $\mathfrak{a}$ over $\boldsymbol{C}$. vii) 1 will often denote the identity element of the group.

CONTENTS.

§ 1. Preliminaries from the theory of Lie groups and hermitian symmetric spaces.

§ 2. Canonical factor of automorphy and automorphic forms of type ρ.

§ 3. The vector bundles $E_{\rho \circ J}$ over $Y = \Gamma \backslash X$ and the fundamental bundle diagram.

§ 4. Line bundles over Y_T.

§ 5. Chern classes and positive line bundles.

§ 6. Generalization of Hirzebruch's proportionality principle.

§ 7. Vanishing theorems for the cohomology groups $H^q(Y, \boldsymbol{E}_{\rho\circ J})$.
§ 8. The proof of Proposition 7.
§ 9. The case of Siegel's symplectic space.
§ 10. Final remarks.
References.

§ 1. Preliminaries from the theory of Lie groups and hermitian symmetric spaces.

1. This section will be devoted to a brief summary of the properties of symmetric hermitian manifolds which should become necessary in the subsequent sections.[5]

Let X denote a symmetric bounded domain of N-complex dimensions, and X^u the compact symmetric hermitian manifold which is dual to X in the sense of E. Cartan (cf. [6; 10]). The group-theoretical descriptions of X and X^u are as follows: We denote by G^c the *simply-connected* covering group of the connected holomorphic automorphisms group of X^u; G^c is a connected complex semi-simple Lie group, and X^u is represented as the complex coset space G^c/U where U is the connected closed complex Lie subgroup of G^c consisting of all element of G^c which fix a point (which we denote by $x_0{}^u$). A compact form G^u of G^c is also *simply-connected*, and acts on X^u as the transitive isometric transformations group; therefore X^u is also written as G^u/K, $K = G^u \cap U$. Further we know that there is an involutive automorphism ϵ of G^u such that K coincides with the subgroup of G^u consisting of all $g \in G^u$ fixed by ϵ. Denote now by $\mathfrak{g}^u$ (resp. $\mathfrak{g}^C, \mathfrak{k}$) the Lie algebra of G^u (resp. of G^c, of K), then the differential $\dot{\epsilon}$ of ϵ yields the decomposition:

$$\mathfrak{g}^u = \mathfrak{k} \dotplus \mathfrak{m} \quad (\text{hence } \mathfrak{g}^c = \mathfrak{k}^c \dotplus \mathfrak{m}^c).$$

where $\mathfrak{m}$ denotes the (-1)-eigen space of $\dot{\epsilon}$ in $\mathfrak{g}^u$; we know further that

$$[\mathfrak{k}, \mathfrak{m}] = \mathfrak{m}, \qquad [\mathfrak{m}, \mathfrak{m}] = \mathfrak{k}.$$

A non-compact real form $\mathfrak{g}$ of $\mathfrak{g}^c$ will be now defined by

$$\mathfrak{g} = \mathfrak{k} \dotplus \sqrt{-1} \cdot \mathfrak{m} \quad (\subset \mathfrak{g}^c);$$

then the real subalgebra $\mathfrak{g}$ of $\mathfrak{g}^c$ generates in G^c the connected real semi-simple Lie subgroup G whose center is finite and simple components are all non-compact. K is then a maximal compact subgroup of G and it is known that

---

[5] For the theory of semi-simple Lie algebras used in this paper, we refer to [13].

$$G\cap G^u=K,\qquad G\cap U=G^u\cap U=K,$$

and that G/K is identified with X. Therefore, if we define the mapping j of X into X^u by

$$j\colon gK\to gU\quad (g\in G),$$

then j yields an *injection* of X into X^u which is compatible with the actions of G on X and on X^u. Thus X will be endowed with the G-invariant complex structure from that of the open submanifold $j(X)$ of X^u; this coincides with the complex structure as a symmetric bounded domain (see, for these arguments, Borel [1]).[6]

2. Now taking and fixing a maximal toral subgroup T of K (then it is also a maximal toral subgroup of G^u), we denote by $\mathfrak{t}$ the Lie algebra of T; the complexified form $\mathfrak{t}^c$ of $\mathfrak{t}$ gives a Cartan subalgebra of $\mathfrak{k}^c$ and also of $\mathfrak{g}^c$. In the sequel, we consider the Cartan decomposition of $\mathfrak{g}^c$ (resp. $\mathfrak{k}^c$) with respect to $\mathfrak{t}$ and the corresponding root system; for a root $\mathring{\alpha}$, the non-zero element $e_\alpha\in\mathfrak{g}^c$ such that $[h,e_\alpha]=\mathring{\alpha}(h)e_\alpha$ $(h\in t^c)$ and $\Phi(e_\alpha,e_{-\alpha})=1$ (Φ denotes the Killing form of $\mathfrak{g}^c$) is called the root vector corresponding to $\mathring{\alpha}$. We will now identify $\mathfrak{m}$ with the tangent space of X^u at the point $x_0{}^u$; then the tensor field I of type (1.1) defining the complex structure of X^u induces a linear transformation I_0 of $\mathfrak{m}$ which commutes with the adjoint actions of K on $\mathfrak{m}$ and $I_0{}^2=-1$. Then we have the decomposition

$$\mathfrak{m}^c=\mathfrak{n}^+\dotplus\mathfrak{n}^-,\quad \mathfrak{n}^+\cap\mathfrak{n}^-=0,$$

where $\mathfrak{n}^+$ (resp. $\mathfrak{n}^-$) denotes the eigen-space of I_0 corresponding to the eigenvalue $\sqrt{-1}$ (resp. $-\sqrt{-1}$) (In other words, $\mathfrak{n}^+$ (resp. $\mathfrak{n}^-$) consists of all complex tangent vectors at $x_0{}^u$ $(\in\mathfrak{m}^c)$ which annihilate all local holomorphic functions at $x_0{}^u$)). Since $\mathrm{Ad}\,K\cdot\mathfrak{m}\subset\mathfrak{m}$, we see also

$$\mathrm{Ad}\,K\cdot\mathfrak{n}^\pm\subset\mathfrak{n}^\pm,\text{ and hence }[\mathfrak{k}^c,\mathfrak{n}^\pm]\subset\mathfrak{n}^\pm.$$

Therefore we have in particular $[\mathfrak{t}^c,\mathfrak{n}^\pm]\subset\mathfrak{n}^\pm$; from this we infer readily that $\mathfrak{n}^+$ (resp $\mathfrak{n}^-$) has to be spanned by some root vectors. We know furthermore that $e_\alpha\in\mathfrak{n}^+$ if and only if $e_{-\alpha}\in\mathfrak{n}^-$, and that we may choose a linear order for the roots such that if $e_\alpha\in\mathfrak{n}^+$ (resp. $e_\alpha\in\mathfrak{n}^-$) then $\mathring{\alpha}$ is positive (resp. negative); we will fix such a linear order once for all. It follows then that there is uniquely determined the simple roots system $\{\mathring{\alpha}_1,\mathring{\alpha}_2,\cdots,\mathring{\alpha}_l\}$ $(l=\dim T)$;

---

[6] The results of j stated in this section shall be, in reality, proved through the roots-arguments; which is given in the section 2, 3 and 4.

moreover there exists a set $\mathcal{S}$ consisting of some simple roots such that the derived algebra $[\mathfrak{k}^c, \mathfrak{k}^c]$ of $\mathfrak{k}^c$ is the semi-simple Lie subalgebra of $\mathfrak{g}^c$ generated by all root vectors e_{α_i} corresponding to $\mathring{\alpha}_i \in \mathcal{S}$. A (positive) root $\mathring{\alpha}$ such that $e_\alpha \in \mathfrak{n}^+$ (i.e. $\mathring{\alpha}$ is not to be expressed as an integral linear combination of $\mathring{\alpha}_i \in \mathcal{S}$) is called a *complementary* (*positive*) *root.*

Let $\mathfrak{u}$ be the Lie algebra of U; then we know that $\mathfrak{n}^- \subset \mathfrak{u}$. Further we can show that $\mathfrak{n}^+$ (resp. $\mathfrak{n}^-$) is an abelian subalgebra of $\mathfrak{g}^c$, and that

$$\mathfrak{g}^c = \mathfrak{u} \dot{+} \mathfrak{n}^+, \quad \mathfrak{u} = \mathfrak{k}^c \dot{+} \mathfrak{n}^-;$$

hence we have $U = K^c N^-$, (the semi-direct product; cf. 4). Now, for the Cartan decomposition of $\mathfrak{g}^c$ (resp. $\mathfrak{k}^c$) with respect to $\mathfrak{t}^c$, we put

$$\mathfrak{n}_0{}^+ = \sum_{\alpha>0} \boldsymbol{C} e_\alpha, \quad \mathfrak{n}_0{}^- = \sum_{\alpha<0} \boldsymbol{C} e_\alpha, \quad \mathfrak{p}^c = \sum_{e_\alpha \in \mathfrak{k}^c} \boldsymbol{C} \cdot e_\alpha.$$

Then $\mathfrak{g}^c = \mathfrak{t}^c \dot{+} \mathfrak{n}_0{}^+ \dot{+} \mathfrak{n}_0{}^-$, $\mathfrak{k}^c = \mathfrak{t}^c \dot{+} \mathfrak{p}^c$, and $\mathfrak{n}_0{}^\pm$ are maximal nilpotent subalgebras of $\mathfrak{g}^c$; moreover we have $\mathfrak{k} = \mathfrak{t} \dot{+} \mathfrak{p}$ for $\mathfrak{p} = \mathfrak{p}^c \cap \mathfrak{k}([\mathfrak{t}, \mathfrak{p}] \subset \mathfrak{p})$. We define here the complex solvable subalgebras $\mathfrak{u}_0$, $\mathfrak{u}_1$ of $\mathfrak{g}^c$ respectively by

$$\mathfrak{u}_0 = \mathfrak{t}^c + \mathfrak{n}_0{}^-, \qquad \mathfrak{u}_1 = \mathfrak{u}_0 \cap \mathfrak{k}^c;$$

U_0 (resp. U_1, $N_0{}^\pm$) denotes the connected complex subgroup of G^c generated by $\mathfrak{u}_0$ (resp. $\mathfrak{u}_1$, $\mathfrak{n}_0{}^\pm$) $(U_1 \subset U_0 \subset U)$. Then it is easy to verify that

$$K \cap U_1 = T, \quad G \cap U_0 = G^* \cap U_0 = T, \quad U_0 = T^c N_0{}^-.$$

3. On the other hand, (,) will denote the non-degenerate inner product induced on $\mathfrak{t}^c$ by the Killing form of $\mathfrak{g}^c$: For any linear form $\mathring{\sigma}$ on $\mathfrak{t}^c$, there exists an element $h'_\sigma \in \mathfrak{t}^c$ such that $\mathring{\sigma}(h) = (h'_\sigma, h)$ for all $h \in \mathfrak{t}^c$. We set, for two linear forms $\mathring{\sigma}$, τ, $(\mathring{\sigma}, \tau) = (h'_\sigma, h'_\tau)$; then $(\mathring{\sigma}, \mathring{\tau})$ is a non-degenerate bilinear form on the dual space $(\mathfrak{t}^c)^*$ of $\mathfrak{t}^c$. Denote now by h_σ (for $\mathring{\sigma} \in (\mathfrak{t}^c)^*$) the element of $\mathfrak{t}^c$ such that $\mathring{\tau}(h_\sigma) = 2(\mathring{\tau}, \mathring{\sigma})/(\mathring{\sigma}, \mathring{\sigma})$ for any $\mathring{\tau} \in (\mathfrak{t}^c)^*$. For any simple root $\mathring{\alpha}_i$ $(1 \leqq i \leqq l)$, in particular, we put $h_{\alpha_i} = h_i$. We know that $[e_\alpha, e_{-\alpha}] = h'_\alpha = \dfrac{(\alpha \cdot \alpha)}{2} h_\alpha$ for every root $\mathring{\alpha}$. A *weight* $\mathring{\sigma}$ on $\mathfrak{k}^c$ is by definition a linear form on $\mathfrak{t}^c$ such that $\mathring{\sigma}(h_i)$ $(1 \leqq i \leqq l)$ are all integers, especially the weights $\breve{\Lambda}_i$ $(1 \leqq i \leqq l)$ such that $\breve{\Lambda}_i(h_j) = \delta_{ij}$ are called the *fundamental dominant weights.* The weights $2\sum_{i=1}^{l} \mathring{\Lambda}_i$ is equal to the total sum of positive roots; we denote it by $\mathring{\delta}$, while the total sum of complementary positive roots will be denoted by $\dot{\delta}_K$. We set further $\delta_M = 2 \sum_{\alpha_i \in \mathcal{S}} \mathring{\Lambda}_i$ (this is not equal to the total sum of positive roots belonging to $[\mathfrak{k}^c, \mathfrak{k}^c]$!). To every weight $\mathring{\sigma}$ corresponds naturally the *character* $\sigma \in \mathrm{Hom}(T^o, \boldsymbol{C}^*)$ such that $\mathring{\sigma}$ is the differential of σ, since G^o is simply-connected.

4. Finally we shall recall some results of Harish-Chandra [8]: G^o contains $N^+K^oN^-$ as an open set, and the mapping

$$(x, y, z) \to xyz \quad (x \in N^+, y \in K^o, z \in N^-)$$

is an injective holomorphic one of the complex manifold $N^+ \times K^o \times N^-$ into G^o; hence we have, in particular, $K^o \cap N^- = \{1\}$. Further, it holds:

$$G \subset N^+K^oN^- = N^+U.$$

Hence we see that GU is an open set of N^+U and N^+U is that of G^o; factoring by U from the right these sets, we get, using $G \cap U = K$ (cf. 1) and $N^+ \cap U = \{1\}$, the inclusions:

$$X = G/K \subset N^+ \subset G^o/U = X^*.$$

We shall denote by j_1 (resp. j_2) the first (resp. the second) inclusion mapping; they are biholomorphic and the composition $j_2 \circ j_1$ coincides with Borel's imbedding j of X into X^* (cf. [1]), while the mapping $\log \circ j_1 : X \to \mathfrak{n}^+$ is no other than Harish-Chandra's realization of X as a bounded domain in the complex cartesian space $\mathfrak{n}^+$ [8]. Now we show

LEMMA 1. *$G \cdot j_1(X) \subset N^+K^oN^-$, and j_1 is compatible with the actions of G.*

Proof. Let $g \in G$ and $x = g' \cdot x_0 \in X$. Then $g' \equiv j_1(x) \bmod U$, and so $gg' \equiv g \cdot j_1(x) \bmod U$. While $gg' \equiv j_1(gx) \bmod U$, hence it holds that $g \cdot j_1(x) \equiv j_1(gx) \bmod U$. This proves the lemma.

§ 2. Canonical factor of automorphy and automorphic forms of type ρ.

5. By Lemma 1 in the preceding section, we can write for every $g \in G$ and $x \in X$ as follows:

$$g \cdot j_1(x) \equiv j_1(g \cdot x) J(g, x) \mod N^-,$$

where $j_1(gx)$ and $J(g, x)$ denote the N^+-component and the K^o-component of $g \cdot j_1(x)$ respectively. We obtain thus the mapping J of $G \times X$ into K^o which has the following properties:

i) the mapping $x \to J(g, x)$ is a holomorphic one of X into K^o for any fixed $g \in G$.

ii) $J(gg', x) = J(g, g'x)J(g', x)$ for every $g, g' \in G$ and $x \in X$.

iii) $J(g, x_0)$ is the K^c-component of g in $N^+K^cN^-$, and $J(k, x) = k$ for every $k \in K^c$ and $x \in X$.

iv) $J(1, x) = 1$ for every $x \in X$; we have thereby $J(g, x)^{-1} = J(g^{-1}, gx)$ for every $g \in G$ and $x \in X$.

Definition 1. The mapping $J: G \times X \to K^c$ introduced above is called the *canonical factor of automorphy* with respect to the symmetric bounded domain $X = G/K$.

Given any homomorphism $\rho \in \mathrm{Hom}(K^c. GL(m, \boldsymbol{C}))$, we will now define the mapping $J_\rho = \rho \circ J: G \times X \to GL(m, \boldsymbol{C})$ by

$$J_\rho(g, x) = \rho(J(g, x))\,; \text{ for } g \in G, x \in X.$$

J_ρ will be called the *automorphic factor of type* ρ. From the properties of J, we deduce the following:

i)′ J_ρ is holomorphic with respect to the variable $x \in X$.

ii)′ $J_\rho(gg', x) = J_\rho(g, g'x)J_\rho(g', x)$ for $g, g' \in G$ and $x \in X$.

iii)′ $J_\rho(k, x) = \rho(k)$ for $k \in K^c$ and $x \in X$.

iv)′ $J_\rho(1, x) = 1$ for $x \in X$, and $J_\rho(g, x)^{-1} = J_\rho(g^{-1}, gx)$ for $g, g' \in G$ and $x \in X$.

Let now Γ be a discrete subgroup of G which satisfies the following two conditions:

(a) The right coset space $\Gamma\backslash G$ is compact; what is the same, the quotient space $\Gamma\backslash X$ is compact.

(b) Every element $\gamma \in \Gamma$ different from the identity 1, has no-fixed point in X; this implies, as is well known, the quotient space $\Gamma\backslash X$ has the structure of a complex analytic manifold without singularities.

We shall assume throughout this paper these two conditions (a) and (b) for Γ; the quotient manifold $\Gamma\backslash X$ will be denoted by Y (the projection of X onto Y will be denoted by φ). Y is an algebraic manifold by a theorem of Kodaira [15].

Definition 2. A holomorphic mapping f of X into $\boldsymbol{C}^m$ is called an *automorphic form of type* ρ with respect to Γ, if the following conditions are satisfied.

$$f(\gamma x) = \rho(J(\gamma, x))f(x),$$

for $\gamma \in \Gamma$ and $x \in X$.

To make clear the meaning of the above definitions, we will now introduce the holomorphic vector bundle $E_{\rho \circ J}$ over Y as follows: We assume henceforth that $\rho(Z) = \{1\}$ for the finite center Z of G^o and define the action of $g \in G$ on the product manifold $X \times \boldsymbol{C}^m$ by

$$g \circ (x, \xi) = (g \cdot x, \rho(J(g, x))\xi)$$

for $x \in X$ and $\xi \in \boldsymbol{C}^m$, and set $E_{\rho \circ J}$ as the quotient $\Gamma \backslash (X \times \boldsymbol{C}^m)$ by these actions. $E_{\rho \circ J}$ is clearly a complex manifold, with the structure of a holomorphic vector bundle over $Y = \Gamma \backslash X$ with the fibre $\boldsymbol{C}^m$ and the group $GL(m, \boldsymbol{C})$. Then we see readily that an automorphic form of type ρ is naturally identified with a holomorphic cross section of $E_{\rho \circ J}$, and conversely. Thus, the complex vector space of all automorphic forms of type ρ will be denoted by $H^0(Y, E_{\rho \circ J})$.

6. Following to Bott [5], we recall here the concept of homogeneous vector bundles over X^u.[7]

The connected complex Lie group $U = K^c \cdot N^-$ being the semidirect product of K^c and the normal closed subgroup N^- (§1, 4), every $\rho \in \operatorname{Hom}(K^c, GL(m, \boldsymbol{C}))$ can be extended naturally to the holomorphic representation of U over $\boldsymbol{C}^m$, which we will denote also by ρ: $\rho \in \operatorname{Hom}(U, GL(m, \boldsymbol{C}))$. Conversely for every *completely reducible* holomorphic representation $\rho \in \operatorname{Hom}(U, GL(m, \boldsymbol{C}))$, we can show that $\rho(N^-) = \{1\}$; so that we may regard as $\rho \in \operatorname{Hom}(K^c, GL(m, \boldsymbol{C}))$. Thus, $\operatorname{Hom}(K^c, GL(m, \boldsymbol{C}))$ will be identified from now on with a subset of $\operatorname{Hom}(U, GL(m, \boldsymbol{C}))$ in this sense. Now, for any $\rho \in \operatorname{Hom}(K^c, GL(m, \boldsymbol{C}))$, we set $E_\rho{}^u$ as the quotient manifold $G \times_U \boldsymbol{C}^m$ of $G^c \times \boldsymbol{C}^m$ by U under the actions such that

$$u \circ (g, \xi) = (gu^{-1}, \rho(u)\xi)$$

for $g \in G^c$, $\xi \in \boldsymbol{C}^m$ and $u \in U$. Then $E_\rho{}^u$ has the structure a holomorphic vector bundle over X^u with the fibre $\boldsymbol{C}^m$; we call $E_\rho{}^u$ the *homogeneous vector bundle* over X^u with respect to ρ. In the sequel, we will denote by $[g, \xi]_\rho$ the element of $E_\rho{}^u$ which is the natural projection of $(g, \xi) \in G^c \times \boldsymbol{C}^m$. It is to be noted that G^c acts naturally on $E_\rho{}^u$ from the left as bundle automorphisms; in fact, each $g' \in G^c$ acts on $E_\rho{}^u$ by $g' \circ [g, \xi]_\rho = [g'g, \xi]_\rho$.

7. Let us now consider the induced bundle $j^*E_\rho{}^u$ over X of $E_\rho{}^u$ by the mapping j, then, j being compatible with the actions of G, G act on $j^*E_\rho{}^u$ from the left as bundle automorphisms; in particular Γ act on $j^*E_\rho{}^u$. Therefore the quotient manifold $\Gamma \backslash j^*E_\rho{}^u$ is meaningful. It clearly constitutes a holomorphic vector bundle over $Y=\Gamma\backslash X$ with fibre $\boldsymbol{C}^m$. What is now to be proved is

---

[7] See also [12, Chap. I].

PROPOSITION 1. *The vector bundle $E_{\rho\circ J}$ is holomorphically isomorphic to $\Gamma\backslash j^*E_\rho{}^u$.*

Proof. The induced bundle $j_2{}^*E_\rho{}^u$ over N^+ may be regarded to consist of all the elements $[g,\xi]_\rho$ such that $g\in N^+U$ and $\xi\in\boldsymbol{C}^m$. So that, denoting by g_{N^+} (resp. g_U) the N^+-component (resp. U-component) of $g\in N^+U$, it corresponds to such $[g,\xi]_\rho$ the element $(g_{N^+},\rho(g_U)\xi)\in N^+\times\boldsymbol{C}^m$ definitely. This yields an isomorphism of the bundle $j_2{}^*E_\rho{}^u$ onto the trivial bundle $N^+\times\boldsymbol{C}^m$ over N^+ with the fibre $\boldsymbol{C}^m$. From this we infer immediately that the induced bundle $j^*E_\rho{}^u=j_1{}^*(j_2{}^*E_\rho{}^u)$ over X is considered, as the set, to consist of $[g,\xi]_\rho$; $g\in GU$ and $\xi\in\boldsymbol{C}^m$, and that it is isomorphic to the trivial bundle $X\times\boldsymbol{C}^m$ *via* the bundle isomorphism:

$$[g,\xi]_\rho\to(gU,\rho(g_U)\xi),$$

where g_U denotes the U-component of $g\in GU$ in the sense of the inclusion $GU\subset N^+U$. Now, the action of $\gamma\in\Gamma$ on $X\times\boldsymbol{C}^m$ *via* the above bundle isomorphism is given by the formula:

$$\gamma\circ(x,\eta)=(\gamma x,\rho(J(\gamma,x))\eta)$$

for $(x,\eta)\in X\times\boldsymbol{C}^m$. In fact, we may write $g=g'u$; $g'\in G$, $u\in U$ and $(\gamma g')_U=J(\gamma g',x)n'=J(\gamma,g'\cdot x)J(g',x)n'$ (**5**, ii)) for some $n'\in N^-$, and so we have

$$\begin{aligned}\rho((\gamma g')_U)&=\rho(J(\gamma,g'\cdot x_0))\cdot\rho(J(g',x_0))\\&=\rho(J(\gamma,x))\cdot\rho(g'_U);\end{aligned}$$

therefore we have $\rho((\gamma g)_U)=\rho(J(\gamma,x))\cdot\rho(g_U)$ for $x=g'\cdot x_0=g\cdot x_0$. Hence, it follows that $\Gamma\backslash j^*\cdot E_\rho{}^u\cong E_{\rho\circ J}$ *via* the isomorphism: $j^*E_\rho{}^u\cong X\times\boldsymbol{C}^m$. Our proposition is thus proved.

As the special cases of Proposition 1, we shall now concern ourselves with the tangent bundle $\mathfrak{T}_Y$ and the canonical bundle K_Y of Y ($\mathfrak{T}_Y$ will be denoted simply by $\mathfrak{T}$). Denote by $\mathfrak{T}^u$ and K^u the tangent bundle and the canonical bundle of X^u respectively. Then $j^*\mathfrak{T}^u$ (resp. j^*K^u) gives clearly the tangent bundle (resp. the canonical bundle) of X; this implies that $\Gamma\backslash j^*\mathfrak{T}^u\cong\mathfrak{T}$ and that $\Gamma\backslash j^*\cdot K^u\cong K_Y$.

On the other hand, if we denote by $(\mathrm{Ad}_K,\mathfrak{n}^+)$ the adjoint representation of K over $\mathfrak{n}^+$ (this is meaningful since $[\mathfrak{k}^c,\mathfrak{n}^+]\subset\mathfrak{n}^+$; see § 1, 2), then $\mathfrak{T}^u$ is the homogeneous vector bundle over X^u with respect to Ad_K: $\mathfrak{T}^u\cong E_{\mathrm{Ad}_K}{}^u$

(cf. [5]). Alternatively it holds that $K^* \cong E_{\delta_K^{-1*}}$, where $\delta_K \in \mathrm{Hom}(K^c, \boldsymbol{C}^*)$ is the character whose differential $\mathring{\delta}_K$ is the sum of all positive complementary roots (cf. [5]). Hence, from Proposition 1 one concludes:

COROLLARY. $\mathfrak{S} = E_{\mathrm{Ad}_K \circ J}$, *and* $K_Y = E_{\delta_K^{-1} \circ J}$.

§ 3. Fundamental bundle diagram.

8. We recall that T is a fixed maximal toral subgroup of K (also of G^*); for this T we now introduce the following manifolds:

$$X_T = G/T, \quad Y_T = \Gamma\backslash X_T = \Gamma\backslash G/T, \quad X_T{}^* = G^*/T.$$

The last manifold $X_T{}^*$, the so-called *flag manifold* of G^*, is represented as the complex coset space G^c/U_0, where U is a maximal solvable subgroup of G^c such that $U \supset U_0 \supset T^c$ (cf. § 1, 2). If we define the mapping j_T of X_T into $X_T{}^*$ by

$$j_T(gT) = gU_0, \qquad g \in G,$$

then j_T is one-to-one and homeomorphic by $G \cap U_0 = T$ (§ 1, 2). Therefore, X_T has the G-invariant structure induced by j_T from that of the open submanifold $j_T(X_T)$ of $X_T{}^*$. Further, Γ acting on X_T such that every $\gamma \in \Gamma$ $(\gamma \neq 1)$ has no fixed point, Y_T is also a complex manifold; the projection of X_T onto Y_T will be denoted by φ_T.

Now we shall consider three mappings:

$$\begin{aligned} \varpi &: Y_T \to Y, \\ \pi &: X_T \to X, \\ \phi &: X_T{}^* \to X, \end{aligned}$$

which are defined respectively by setting $\varpi(\Gamma gT) = \Gamma gK$ $(g \in G)$, $\pi(gT) = gK$ $(g \in G)$ and $\phi(gT) = gK$ $(g \in G^*)$. Then, they are surjective and clearly constitute the following commutative diagram.

$$\begin{tikzcd}
Y_T \arrow[d, "\varpi"'] & X_T \arrow[l, "\varphi_T"'] \arrow[r, "j_T"] \arrow[d, "\pi"] & X_T{}^* \arrow[d, "\phi"'] \\
Y & X \arrow[l, "\varphi"'] \arrow[r, "j"] & X^*
\end{tikzcd}$$

We set moreover $M = K/T$, then M is the flag manifold of K and is also identified with the complex coset space K^c/U_1; where U_1 is a maximal solvable subgroup of K^c such that $U_0 \supset U_1 \supset T^c$ (§ 1, 2). The holomorphic mapping $\phi : X_T{}^* \to X^*$ clearly makes a holomorphic fibre bundle $X_T{}^*(X^*, M, K^c, \phi)$

with the standard fibre M and the structure group K^c. Therefore the mapping $\pi: X_T \rightarrow X$ also becomes a holomorphic fibre bundle with fibre M as the induced bundle of the bundle $X_T^u(X^u, M, K^c, \phi)$ by j; j_T being the bundle mapping of the former into the latter. Finally, by the commutativity $\varpi \circ \phi_T = \phi \circ \pi$, we infer easily that $\varpi: Y_T \rightarrow Y$ constitutes naturally the holomorphic fibre bundle $Y_T(Y, M, K^c, \varpi)$, where Y_T is an algebraic manifold by a theorem of Borel (cf. [10, p. 140]). It is now to be noted that the bundle $X_T(X, M, K^c, \pi)$ is holomorphically trivial, since X is a contractible Stein manifold. In fact, it holds:

LEMMA 2. *The complex manifold X_T can be identified with $X \times M$ by the mapping:*

$$gT \rightarrow (gK, J(g, x_0)U_1)$$

for $g \in G$; under this identification, the left translation by $g \in G$ on X_T is transferred to the action on $X \times M$ such that

$$g \circ (x, z) = (g \cdot x, J(g, x)z)$$

for $x \in X$ and $z \in M$.

Proof. First, our mapping is surjective, since $J(k, x_0) = k$ for any $k \in K$ (§2, iii)). Suppose that $(gK, J(g, x_0)U_1) = (g'K, J(g', x_0)U_1)$ $(g' \in G)$, then $g' = gk$ for some $k \in K$, and so $J(g', x_0) = J(g, x_0)k$. This implies that $k \in U_1 \cap K = T$ (cf. §1, 2) and that $g' \in gT$. Hence our mapping is injective. Finally it is holomorphic since GU_0 is an open set of G and $GU_0/U_0 = X_T$ as the complex manifolds. The rule for actions of Γ on $X \times M$ is obvious. This furnishes the proof.

Furthermore we define here the *canonical* (holomorphic) injections of the complex manifold M into the complex manifolds Y_T, X_T and X_T^u respectively; $\iota: M_T \rightarrow Y_T$, $\iota_T: M_T \rightarrow X_T$ and $\iota_u: M \rightarrow X_T^u$ as follows:

$$\iota(kT) = \Gamma kT, \quad (k \in K),$$
$$\iota_T(kT) = kT, \quad (k \in K),$$
$$\iota_u(kT) = kT, \quad (k \in K).$$

Then we have obviously

$$\iota = \varphi_T \circ \iota_T, \qquad \iota_u = j_T \circ \iota_T.$$

§ 4. Line bundles over Y_T.

9. The present Section 4 will be devoted to the study of the group $H^1(Y_T, \mathfrak{O}^*)$ of all holomorphic line bundles over Y_T; for this purpose, we shall first determine the subgroup $\mathfrak{H}_M$ of $H^1(Y_T, \mathfrak{O}^*)$ which consists of line bundles along the fibres in the bundle $Y_T(Y, M, K^c, \varpi)$. In what follows, we recall that both G^c and G^u are simply-connected (cf. § 1).

First we consider the holomorphic line bundles over $X_T{}^u$ (resp. over M). Now recall that U_0 (resp. U_1) is the semidirect product of T^c and its derived (normal) subgroup (cf. § 1, 2); hence we can identify:

$$\mathrm{Hom}(T^c, \boldsymbol{C}^*) = \mathrm{Hom}(U_0, \boldsymbol{C}^*) = \mathrm{Hom}(U_1, \boldsymbol{C}^*),$$

in a similar way to the case of § 2, 6. Therefore every $\sigma \in \mathrm{Hom}(T^c, \boldsymbol{C}^*)$ will define the homogeneous line bundle $F_\sigma{}^u$ over $X_T{}^u = G^c/U_0$ (resp. $F_\sigma{}^M$ over $M = K^c/U_1 = U/U_0$). In fact $F_\sigma{}^u$ (resp. $F_\sigma{}^M$) is, by definition, the quotient of $G^c \times \boldsymbol{C}$ (resp. $K^c \times \boldsymbol{C}$) under the actions of U_0 (resp. U_1) such that

$$u \circ (g, \xi) = (gu^{-1}, \sigma(u)\xi)$$
$$(\text{resp.}\, u \circ (k, \xi) = (ku^{-1}, \sigma(u)\xi))$$

for $(g, \xi) \in G^c \times \boldsymbol{C}$ and $u \in U_0$ (resp. for $(k, \xi) \in K^c \times \boldsymbol{C}$ and $u \in U_1$); we thus write $F_\sigma{}^u = G^c \times_{U_0} \boldsymbol{C}$ (resp. $F_\sigma{}^M = K^c \times_{U_1} \boldsymbol{C}$). Moreover, the element of $F_\sigma{}^u$ containing (g, ξ) (resp. the element of $F_\sigma{}^M$ containing (k, ξ)) will be denoted by $[g, \xi]_\rho$ (resp. $[k, \xi]_\sigma$). We have then clearly

$$\iota_u{}^* F_\sigma{}^u = F_\sigma{}^M.$$

The mapping $\sigma \rightarrow F_\sigma{}^u$ gives rise to an *isomorphism* of $\mathrm{Hom}(T^c, \boldsymbol{C}^*)$ onto $H^1(X^u, \mathfrak{O}^*)$ since G^c is simply-connected [12, Theorem 1]; while $\sigma \rightarrow F_\sigma{}^M$ is a homomorphism of $\mathrm{Hom}(T^c, \boldsymbol{C}^*)$ into $H^1(M, \mathfrak{O}^*)$ whose kernel consists of σ such that σ is extendable to a character $\sigma \in \mathrm{Hom}(K^c, \boldsymbol{C}^*)$ [12, Lemma 3]. Further we want to show the

LEMMA 3. *The following exact sequence of abelian groups holds:*

$$0 \rightarrow \mathrm{Hom}(K^c, \boldsymbol{C}^*) \rightarrow \mathrm{Hom}(T^c, \boldsymbol{C}^*) \rightarrow H^1(M, \mathfrak{O}^*) \rightarrow 0.$$

Proof. The proof is quite similar to that of Theorem 1 in [12]; however we state it here for the completeness sake. Both T and T^1 ($=$ 1-dimensional real torus) being the maximal compact subgroups of T^c and $\boldsymbol{C}^*$ respectively, we infer easily that $\mathrm{Hom}(T^c, \boldsymbol{C}^*) \cong \mathrm{Hom}(T, T^1)$, while $\mathrm{Hom}(T, T^1) \cong H^1(T, \boldsymbol{Z})$ as was shown in [4, § 10]. We see thus $\mathrm{Hom}(T^c, \boldsymbol{C}^*) = H^1(T, \boldsymbol{Z})$. On the other hand, in the principal bundle $K(M, T)$, the transgression τ will clearly map $H^1(T, \boldsymbol{Z})$ onto $H^2(M, \boldsymbol{Z})$ since the transgression: $H^1(T, \boldsymbol{Z}) \rightarrow H^2(X_T{}^u, \boldsymbol{Z})$ is bijective (cf. [4]) and the restriction map $\iota_u{}^*: H^2(X_T{}^u, \boldsymbol{Z}) \rightarrow H^2(M, \boldsymbol{Z})$ is surjective. Further, the first Betti number of M vanishing, the characteristic homomorphism c yields an isomorphism of $H^1(M,\mathfrak{D}^*)$ onto $H^2(M,\boldsymbol{Z})$ $(=H^2_{1.1}(M,\boldsymbol{Z}))$. As was already shown in [4, Theorem 10.3], it holds the commutative diagram:

$$\begin{tikzcd}
\mathrm{Hom}(T^c,\boldsymbol{C}^*) \arrow[r] \arrow[d, equal] & H^1(M,\mathfrak{D}^*) \arrow[d, "c"] \\
H^1(T,\boldsymbol{Z}) \arrow[r, "\tau"] & H^2(M,\mathfrak{D}^*).
\end{tikzcd}$$

Hene, the original homomorphism must be surjective; this proves the lemma.

Now we define two subgroups $\mathfrak{X}_K$ and $\mathfrak{X}_M$ of $\mathrm{Hom}(T^c,\boldsymbol{C}^*)$ by setting:

$$\mathfrak{X}_K=\{\lambda\in\mathrm{Hom}\ (T^c,\boldsymbol{C}^*)\mid(\mathring{\lambda},\mathring{\alpha}_i)=0 \text{ for all } \mathring{\alpha}_i\in\mathscr{S}\ {}^{8}\}.$$

$$\mathfrak{X}_M=\{\mu\in\mathrm{Hom}.(T^c,\boldsymbol{C}^*)\mid(\mathring{\mu},\mathring{\alpha}_j)=0 \text{ for all } \mathring{\alpha}_j\notin\mathscr{S}\}.$$

As is readily observed, it holds that $\mathfrak{X}_K\cong\mathrm{Hom}(K^c,\boldsymbol{C}^*)$ *via* the restriction mapping in Lemma 3 (we identify these two groups!) and that $\mathrm{Hom}(T^c,\boldsymbol{C}^*)$ is the direct product of $\mathfrak{X}_K$ and $\mathfrak{X}_M$; we can thus write every $\sigma\in\mathrm{Hom}(T^c,\boldsymbol{C}^*)$ as $\sigma=(\lambda,\mu)$, $\lambda\in\mathfrak{X}_K$, $\mu\in\mathfrak{X}_M$. So that, $\mu\to F_\mu{}^M$ yields the isomorphism:

$$\mathfrak{X}_M\cong H^1(M,\mathfrak{D}^*).$$

From these arguments, it follows now the corollary:

COROLLARY. *We have the commutative diagram of abelian groups (the extensions are splittable):*

$$\begin{tikzcd}
0 \arrow[r] & \mathfrak{X}_K \arrow[r] \arrow[d] & \mathrm{Hom}(T^c,\boldsymbol{C}^*) \arrow[r] \arrow[d] & \mathfrak{X}_M \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & H^1(X^u,\mathfrak{D}^*) \arrow[r, "\phi^*"] & H^1(X_T{}^u,\mathfrak{D}^*) \arrow[r, "\iota_u{}^*"] & H^1(M,\mathfrak{D}^*) \arrow[r] & 0,
\end{tikzcd}$$

where the vertical mappings are all isomorphisms.

10. On the other hand, G acts as bundle automorphisms on the induced bundle $j_T{}^*F_\sigma{}^u$ over X_T; therefore we can define the holomorphic line bundle F_σ over Y_T as the quotient manifold of $j_T{}^*F_\sigma{}^u$ by Γ $(\subset G)$:

$$F_\sigma=\Gamma\backslash j_T{}^*F_\sigma{}^u.$$

We see then readily that

$$\iota^*F_\sigma=F_\sigma{}^M.$$

Define now the subgroup $\mathfrak{H}_M$ of $H^1(Y_T,\mathfrak{D}^*)$ by setting;

---

[8] $\mathscr{S}$ is the simple roots system of $[\mathfrak{k}^c,\mathfrak{k}^c]$; see §1, 2.

$$\mathfrak{H}_M = \{F_\mu \mid \mu \in \mathfrak{X}_M\}.$$

Then the restriction mapping $\iota^* : F_\mu \to F_\mu{}^M$ gives the isomorphism by Corollary of Lemma 3:

$$\mathfrak{H}_M \cong H^1(M, \mathfrak{D}^*).$$

We shall now prove:

PROPOSITION 2. *It holds that*

$$H^1(Y_T, \mathfrak{D}^*) = \varpi^* \cdot H^1(Y, \mathfrak{D}^*) \otimes \mathfrak{H}_M,$$

where $\otimes$ *denotes the tensor product of line bundles, and* $\varpi^* H^1(Y, \mathfrak{D}^*) \cap \mathfrak{H}_M$ *contains only the trivial bundle.*

(This proposition is not needed for the subsequent sections, but it will be necessary in the forthcoming paper.)

Proof. First suppose that $F_\mu \in \varpi^* H^1(Y, \mathfrak{D}^*) \cap \mathfrak{H}_M$, then $\iota^* \cdot F_\mu = F_\mu{}^M$ is clearly the trivial bundle; hence $\mu = 1$ by Cor. of Lemma 3. Therefore F_μ is also trivial.

Next we shall study the second integral cohomology group $H^2(Y_T, \boldsymbol{Z})$ of Y_T; for this sake we consider the following diagram of bundles:

$$\begin{tikzcd}
\Gamma\backslash G \arrow[r, "\varpi_T"] \arrow[dr, "\varpi_K"'] & Y_T \arrow[d, "\varpi"] \\
 & Y,
\end{tikzcd}$$

where $\varpi_T : \Gamma\backslash G \to Y_T$ is given by $\varpi_T(\Gamma g) = \Gamma g T$ and $\varpi_K : \Gamma\backslash G \to Y$ by $\varpi_K(\Gamma g) = \Gamma g K$ respectively; the latter clearly makes the differentiable principal K-bundle over Y with the bundle $Y_T(Y, M, K, \varpi)$ as its associated one. While, $K/T = M$ being the flag manifold, we can apply a result of A. Borel [2, Cor. de Proposition 18.3] to our case. In fact, from this and from that $H^1(M, \boldsymbol{Z}) = 0$, we infer that

$$H^2(Y_T, \boldsymbol{Z}) = \varpi^* H^2(Y, \boldsymbol{Z}) \dotplus c(\mathfrak{H}_M),$$

where $c(\mathfrak{H}_M)$ denotes the subgroup of $H^2(Y_T, \boldsymbol{Z})$ consisting of the Chern classes $c(F_\mu)$ of line bundles $F_\mu \in \mathfrak{H}_M$:

$$c(\mathfrak{H}_M) = \{c(F_\mu) \mid F_\mu \in \mathfrak{H}_M\}.$$

We note here that $H^2(M, \boldsymbol{Z})$ consists of all the Chern classes of

$$F_\mu{}^M \in H^1(M, \mathfrak{D}^*) : H^2(M, \boldsymbol{Z}) = \{c(F_\mu{}^M) \mid \mu \in \mathrm{X}_M\}$$

(cf. [4], [12]), and that $c(\mathfrak{H}_M)$ is isomorphic *via* ι^* to $H^2(M;\boldsymbol{Z})$, since $\iota^* F_\mu = F_\mu{}^M$. Denote now by $H^2{}_{1.1}(Y_T,\boldsymbol{Z})$ (resp. $H^2{}_{1.1}(Y,\boldsymbol{Z})$) the subgroup of $H^2(Y_T,\boldsymbol{Z})$ (resp. $H^2(Y,\boldsymbol{Z})$) consisting of all 2-cohomology classes of type (1.1); that is to say, the inverse image of $H^2{}_{1.1}(Y,\boldsymbol{C})$ (resp. $H^2{}_{1.1}(Y,\boldsymbol{C})$) by the natural homomorphism $H^2(Y,\boldsymbol{Z}) \to H^2(Y,\boldsymbol{C})$ (resp. $H^2(Y,\boldsymbol{Z}) \to H^2(Y,\boldsymbol{C})$), where $H^2{}_{1.1}(Y_T,\mathbf{C})$ (resp. $H^2{}_{1.1}(Y,\boldsymbol{C})$) denotes all the cohomology classes containing closed 2-forms of type (1.1).

Then we see readily that

$$H^2{}_{1.1}(Y_T,\boldsymbol{Z}) = \varpi^* H^2{}_{1.1}(Y,\boldsymbol{Z}) \dotplus c(\mathfrak{H}_M),$$

by the fact that both ϖ and ι are holomorphic. Therefore, the theorem of Lefschetz-Hodge [9, p. 125] implies that

$$c(H^1(Y_T,\mathfrak{O}^*)) = c(\varpi^* H^1(Y,\mathfrak{O}^*)) \dotplus c(\mathfrak{H}_M);$$

c denoting the characteristic homomorphisms. Hence we see that

$$H^1(Y_T,\mathfrak{O}^*) \equiv \varpi^* H^1(Y,\mathfrak{O}^*) \otimes \mathfrak{H}_M$$

modulo the Picard variety $\mathfrak{P}$ of Y_T. Here, by the Picard variety $\mathfrak{P}$ of Y_T, we understand as a subgroup of $H^1(Y_T,\mathfrak{O}^*)$ consisting of certain line bundles associated with characters of the fundamental group $\pi_1(Y_T)$; $\pi_1(Y_T)$ is identified with Γ since $\pi_1(X_T) = 0$. Let $\hat{E}_\sigma \in \mathfrak{P}$ be now associated with $\sigma \in \mathrm{Hom}(\Gamma, \boldsymbol{C}^*)$: $\hat{E}_\sigma = X_T \times_\Gamma \boldsymbol{C}$, where the actions of Γ on $\boldsymbol{C}$ is those of $\sigma(\Gamma)$.[9] We can correspondingly define a line bundle E_σ over Y by the same σ: $E_\sigma = X \times_\Gamma \boldsymbol{C}$ [9]; then we have clearly $\hat{E}_\sigma = \varpi^* \cdot E_\sigma$. Hence we see that $\mathfrak{P} \subset \varpi^* \cdot H^1(Y,\mathfrak{O}^*)$; in other words, it holds that

$$H^1(Y_T,\mathfrak{O}^*) = \varpi^* \cdot H^1(Y,\mathfrak{O}^*) \otimes \mathfrak{H}_M.$$

This completes the proof of our proposition.

11. In the sequel, we shall confine ourselves to the line bundles of the form F_σ; $\sigma = (\lambda, \mu)$ $(\lambda \in \mathfrak{X}_K,\ \mu \in \mathfrak{X}_M)$. Now we put

$$F_{\lambda,\mu} = \varpi^* E_{\lambda \circ J} \otimes F_\mu,$$

then the commutatively of the fundamental bundle diagram (§ **3**) implies that $F_\lambda = \varpi^* E_{\lambda \circ J}$, while $F_\sigma = F_\lambda \otimes F_\mu$; hence we see

$$F_\sigma = F_{\lambda,\mu}.$$

For the later use, we mention here the following lemma.

---

[9] $X_T \times_\Gamma \boldsymbol{C}$ denotes the quotient manifold of $X_T \times \boldsymbol{C}$ by the actions of Γ: $\gamma \circ (x, \xi) = (\gamma x, \sigma(\gamma^{-1})\xi)$ for $\gamma \in \Gamma$, $(x, \xi) \in X_T \times \boldsymbol{C}$. $X \times_\Gamma \boldsymbol{C}$ is also understood in a similar way.

LEMMA 4. *The holomorphic line bundle* F_σ *is identified with*

$$\Gamma\backslash(X\times F_\mu{}^M);$$

the action of Γ *on* $X\times F_\mu{}^M$ *is now given by the formula:*

$$\gamma\circ(x,[k,\xi]_\mu)=(\gamma x,[J(\gamma,x)k,\lambda(J(\gamma,x))\xi]_\mu),$$

for $\gamma\in\Gamma$, $x\in X$ *and* $[k,\xi]_\mu\in F_\mu{}^M$ (cf. § **4**, **9**).

Proof. By Lemma 2, $X_T=X\times M$. The bundle $j_T{}^*F_\sigma{}^u$ is also identified with $X\times F_\mu{}^M$ so that the diagram

$$\begin{tikzcd}
j_T{}^*F_\sigma{}^u \arrow[r] \arrow[d] & X\times F_\mu{}^M \arrow[d] \\
X_T \arrow[r, equal] & X\times M
\end{tikzcd}$$

is commutative. In fact, $j_T{}^*F_\sigma{}^u$ can be regarded as the subset of $F_\sigma{}^u$ consisting of $[g,\xi]_\sigma$; $g\in GU_0$, $\xi\in\boldsymbol{C}$. Thus, denoting by g_{K^c} the K-component of g in the sense of the inclusion $GU_0\subset GU\subset N^+K^cN^-$, we can define the bundle isomorphism of $j_T{}^*F_\sigma{}^u$ onto $X\times F_\mu{}^M$ by

$$[g,\xi]_\sigma\to(gU_0,[g_{K^c},\lambda(g_{K^c})\xi]_\mu).$$

We see, in fact, that for any $u\in U_0$,

$$(gu)_{K^c}=((gu)_U)_{K^c}=(g_U\cdot u)_{K^c}=g_{K^c}\cdot u_{K^c}.$$

On the other hand, we can show that $(\gamma g)_{K^c}=J(\gamma,x)\cdot g_{K^c}$ $(x=gU_0)$ as in the proof of Proposition 1 (§ **2**). Now it is easy to verify that, *via* the above bundle isomorphism, the actions of Γ on $j_T{}^*F_\sigma{}^u$ is given by the formula mentioned in our lemma.

COROLLARY. *The canonical bundle* K_T *of* Y_T *is given by* $K_T=F_{\delta^{-1}}=\Gamma\backslash j_T{}^*F_{\delta^{-1}}{}^u$, *and identified with* $\Gamma\backslash(X\times F_{\delta_M{}^{-1}}{}^M)$.

Proof. We know that the canonical bundle of $X_T{}^u$ is the dual line bundle of $F_\delta{}^u$ (see, for instance, [5]); hence F_δ is isomorphic to $K_T{}^{-1}$. The last part is a special case of Lemma 4.

§ 5. Chern classes and positive line bundles.

12. In this section, we shall discuss the Chern classes of the vector bundle $E_{\rho\circ J}$ and the line bundle F_σ. For this sake, we shall employ an analogue of Borel-Hirzebruch's method in dealing with the Chern classes of the vector bundle $E_{\rho}{}^{u}$ over X^{u} and of the line bundle $F_{\sigma}{}^{u}$ over $X_{T}{}^{u}$ respectively (see [4]; in particular, Theorem 10.3).

Let us now consider the differentiable principal K-bundle over Y: $\Gamma\backslash G(Y, K, \varpi_K)$, where K acts on $\Gamma\backslash G$ from the right, and ϖ_K denotes the natural projection of $\Gamma\backslash G$ onto $Y = \Gamma\backslash G/K$. Alternatively we have the differentiable principal T-bundle over Y_T: $\Gamma\backslash G(Y_T, T, \varpi_T)$, in a similar way. We obtain thus the triangle of bundles, which was already utilized in § **4** (cf. the proof of Proposition 2):

$$\begin{tikzcd}
\Gamma\backslash G \arrow[r, "\varpi_T"] \arrow[d, "\varpi_K"'] & Y_T \arrow[dl, "\varpi"] \\
Y &
\end{tikzcd}$$

For every $\rho \in \mathrm{Hom}(K^c, GL(m, \boldsymbol{C}))$ (resp. $\sigma \in \mathrm{Hom}(T^c, \boldsymbol{C}^*)$), we will denote by ρ_K (resp. by σ_T) the restriction of ρ to K (resp. the restriction of σ to T). We may then assume that $\rho_K(K) \subset U(m)$ and $\sigma_T(T) \subset U(1) = T^1$; so that we have the differentiable $U(m)$-bundle B_ρ over Y as the bundle associated to the bundle $\Gamma\backslash G(Y, K, \varpi_K)$ by ρ_K, and the differentiable $U(1)$-bundle B_σ over Y_T as the associated one to the bundle $\Gamma\backslash G(Y_T, T, \varpi_T)$ by σ_T respectively. Furthermore we will denote the vector bundle associated to the principal bundle B_ρ (resp. to B_σ) with the fibre $\boldsymbol{C}^m$ (resp. with the fibre $\boldsymbol{C}$) by the same letter B_ρ (resp. B_σ).

PROPOSITION 3. i)[10] *$E_{\rho\circ J}$ is differentiably isomorphic to B_ρ; therefore $c_i(E_{\rho\circ J}) = c_i(B_\rho)$ $(1 \leqq i \leqq m)$.*

ii) *F_σ is differentiably isomorphic to B_σ, and so $c(F_\sigma) = c(B_\sigma)$.*

Proof. i) We will define the differentiable mapping ψ of $G \times \boldsymbol{C}^m$ onto $X \times \boldsymbol{C}^m$ such that

$$\psi \circ (g, \xi) = (g \cdot x_0, \rho(J(g, x_0))\xi),$$

for $(g, \xi) \in G \times \boldsymbol{C}^m$. While, the bundle B_ρ is the quotient of $G \times \boldsymbol{C}^m$ with respect to the equivalence relation:

$$(g, \xi) \sim (\gamma g k^{-1}, \rho_K(k)\xi)$$

for $\gamma \in \Gamma$ and $k \in K$, and alternatively $E_{\rho\circ J}$ is the quotient of $X \times \boldsymbol{C}^m$ with respect to the equivalence relation indicated in § **2**. The mapping ψ being clearly compatible with these equivalence relations, it induces the bundle isomorphism of B_ρ onto $E_{\rho\circ J}$.

---

[10] This is due to Y. Matsushima and S. Murakami [16: Part II, Prop. 4.1].

ii) The bundle B_σ is the quotient of $G\times\boldsymbol{C}$ with respect to the equivalence relation:

$$(g,\xi)\sim(\gamma g t^{-1},\sigma_T(t)\xi),$$

for $(g,\xi)\in G\times\boldsymbol{C}$, $\gamma\in\Gamma$ and $t\in T$. While $F_\sigma=F_{\lambda,\mu}$ is the quotient of $X\times F_\mu{}^M$ with respect to the one given in Lemma 4 (§4). Define now the mapping ψ of $G\times\boldsymbol{C}$ onto $X\times F_\mu{}^M$ by putting:

$$\psi\circ(g,\xi)=(g\cdot x_0,[J(g,x_0),\lambda(J(g,x_0))\xi]_\mu),$$

for $(g,\xi)\in G\times\boldsymbol{C}$. Then we can readily verify that ψ is compatible with these equivalence relations. Hence ψ will define the bundle isomorphism of B_σ onto F_σ. Our proposition is now proved.

Owing to this proposition, we may be hereafter concerned with the Chern classes of B_ρ (resp. B_σ), instead of those of $E_{\rho\circ J}$ (resp. F_σ).

13. Recall now that $\mathrm{Hom}(T^0,\boldsymbol{C}^*)$, or the lattice of weights can be identified with $H^1(T,\boldsymbol{Z})$ under the correspondence $\mathring{\sigma}\to\mathring{\sigma}/2\pi\sqrt{-1}$ (see, for detail, [4, §10]) since G^u is simply-connected. We denote by τ^u (resp. by τ) the transgression of the bundle $G^u(X^u,T,\phi_T)$ (resp. $\Gamma\backslash G(Y_T,T,\varpi_T)$). For any given $\rho\in\mathrm{Hom}(K^0,GL(m,\boldsymbol{C}))$ we also denote by $\mathring{\sigma}_1,\cdots,\mathring{\sigma}_m$ the weights of ρ. Then a result of Borel-Hirzebruch [4, Theorem 10.3] implies in our cases the following:

LEMMA 5. i) *For any character* $\sigma\in\mathrm{Hom}(T^0,\boldsymbol{C}^*)$, *it holds that*

$$c(B_\sigma)=-\tau(\mathring{\sigma}/2\pi\sqrt{-1}),\qquad c(B_\sigma{}^u)=-\tau^u(\mathring{\sigma}/2\pi\sqrt{-1}).$$

ii) *The total Chern classes* $c(B_\rho)$, $c(B_\rho{}^u)$ *of the bundles* B_ρ, $B_\rho{}^u$ *are given respectively by*

$$\phi^*c(B_\rho)=\prod_{i=1}^{m}(1-\tau(\mathring{\sigma}_i/2\pi\sqrt{-1}))\in H^*(Y,\boldsymbol{Z}),$$

$$\varpi^*c(B_\rho{}^u)=\prod_{i=1}^{m}(1-\tau^u(\mathring{\sigma}_i/2\pi\sqrt{-1}))\in H^*(X^u,\boldsymbol{Z}).$$

We shall now take out closed differential forms which represent the real classes of the Chern classes appeared above: For $\sigma\in\mathrm{Hom}(T^0,\boldsymbol{C}^*)$ we set

$$\tilde{\eta}_\sigma(x,y)=\frac{\sqrt{-1}}{4\pi}\,\mathring{\sigma}([x,y])$$

for $x,y\in\mathfrak{g}^0$, where $\mathring{\sigma}$ is regarded as a linear form on $\mathfrak{g}^0$ by putting

$$\mathring{\sigma}(h+\sum_{\alpha} c_\alpha e_\alpha)=\mathring{\sigma}(h) \quad (h\in\mathfrak{t}^c, c_\alpha\in \boldsymbol{C})$$

for the Cartan decomposition of $\mathfrak{g}^c$ (cf. § 1). Then we obtain an alternating bilinear form $\tilde{\eta}_\sigma$ on $\mathfrak{g}^c$ which is invariant under Ad T^c. The restriction $\mathring{\sigma}\mid\mathfrak{t}$ being purely-imaginary valued, the restriction $\eta_\sigma=\tilde{\eta}_\sigma\mid\mathfrak{g}$ (resp. $\eta_\sigma{}^u=\tilde{\eta}_\sigma\mid\mathfrak{g}^u$) is an alternating real valued bilinear form on $\mathfrak{g}$ (resp. on $\mathfrak{g}^u$) which is invariant under Ad T. Moreover we see that $\tilde{\eta}_\sigma$ is a 2-cocycle in the sense of relative cohomology of $\mathfrak{g}^c \bmod \mathfrak{t}^c$. This implies that η_σ (resp. $\eta_\sigma{}^u$) induces a G-invariant (resp. G^u-invariant) closed 2-form on $G/T=X_T$ (resp. on $G^u/T=\mathfrak{X}_{T^u}$); these forms shall be denoted by the same letters η_σ, $\eta_\sigma{}^u$ for brevity. Further η_σ naturally defines the closed 2-form $\hat{\eta}_\sigma$ on $\Gamma\backslash X_T=Y_T$. If $\sigma=\lambda\in\mathfrak{X}_K$ in particular, we see similarly that η_λ (resp. $\eta_\lambda{}^u$) represents a G-invariant (resp. G^u-invariant) closed 2-form on X (resp. on X^u) and that η_λ naturally induces the closed 2-form $\hat{\eta}_\sigma$ on $\Gamma\backslash X=Y$. (We often write as $\eta_\sigma{}^u=\eta(B_\sigma{}^u)$, $\hat{\eta}_\sigma=\eta(B_\sigma)$ etc.)

Next, for $\rho\in\operatorname{Hom}(K^c, GL(m,\boldsymbol{C}))$ with the weights $\mathring{\sigma}_1,\cdots,\mathring{\sigma}_m$, we now introduce the alternating $2p$-form $\tilde{\eta}_\rho{}^{(p)}$ on $\mathfrak{g}^c$ by

$$\tilde{\eta}_\rho{}^{(p)}=\sum_{i_1<\cdots<i_p}\tilde{\eta}_{\sigma_{i_1}}\wedge\cdots\wedge\tilde{\eta}_{\sigma_{i_p}} \qquad (1\leqq p\leqq m).$$

Then, $\tilde{\eta}_\rho{}^{(p)}$ is a $2p$-cocycle not only in the sense of relative cohomology of $\mathfrak{g}^c$ *mod* $\mathfrak{t}^c$, but also in the sense of relative cohomology of $\mathfrak{g}^c$ *mod* $\mathfrak{k}^c$. This will be readily verified from the following fact: if we denote by ρ_p the representation of K^c over the alternating tensors of degree p induced by ρ, the traces of linear transformations belonging to $\rho_p([\mathfrak{k}^c,\mathfrak{k}^c])$ vanish. Let $\tilde{\eta}_\rho{}^{(p)}$ (resp. $\eta_\rho{}^{u(p)}$ now denote the restriction of $\tilde{\eta}_\rho$ to $\mathfrak{g}$ (resp $\mathfrak{g}^u$), then we have

$$\eta_\rho{}^{(p)}=\sum_{i_1<\cdots<i_p}\eta_{\sigma_{i_1}}\wedge\cdots\wedge\eta_{\sigma_{i_p}},$$

$$\eta_\rho{}^{u(p)}=\sum_{i_1<\cdots<i_p}\eta_{\sigma_{i_1}}{}^u\wedge\cdots\wedge\eta_{\sigma_{i_p}}{}^u;$$

moreover $\eta_\rho{}^{(p)}$ (resp. $\eta_\rho{}^{u(p)}$ gives rise to a G-invariant (resp. G^u-invariant) closed $2p$-form on X (resp. on X^u). $\eta_\rho{}^{(p)}$ also induces the closed $2p$-form $\hat{\eta}_\rho{}^{(p)}$ on Y. They shall be often denoted respectively by $\eta_\rho{}^{u(p)}=\eta^p(B_\rho{}^u)$ and $\hat{\eta}_\rho{}^{(p)}=\eta^p(B_\rho)$.

Definition 3. The closed 2-form $\eta(B_\sigma{}^u)$ (resp. $\eta(B_\lambda{}^u)$, $\eta(B_\sigma)$, $\eta(B_\lambda)$) is called the *canonical Chern form* of the bundle $B_\sigma{}^u$ over X_{T^u} (resp. $B_\lambda{}^u$ over X^u, B_σ over Y_T, B_λ over Y). While, $\eta^p(B_\rho{}^u)$ (resp. $\eta^p(B_\rho)$) is called the *p-th canonical Chern form* of the bundle $B_\rho{}^u$ (resp. B_ρ).

PROPOSITION 4. *The canonical Chern forms of these* $U(1)$*-bundles are* *all of type* (1.1) *in the respective complex manifolds, and belong to the real Chern classes of the respective* $U(1)$*-bundles.*

From this proposition, combining with Lemma 5, we can deduce immediately the following:

COROLLARY. *The form* $\eta^p(B_\rho)$ *(resp.* $\eta^p(B_\rho{}^u)$*)* $(1 \leqq p \leqq m)$ *is of type* (p,p)*, and belongs to the p-th real Chern class of* B_ρ *(resp. of* $B_\rho{}^u$*).*

14. Now, we shall proceed to the proof of Proposition 4[11]: We denote by τ' the transgression in the principal bundle $G(X_T, T, \pi_T)$, then we can infer easily that $\tau' = \varphi_T{}^* \circ \tau$ in the next diagram.

$$\begin{tikzcd}
H^1(T,\boldsymbol{R}) \arrow[r, "\tau'"] \arrow[dr, "\tau"'] & H^2(X_T,\boldsymbol{R}) \arrow[d, "\varphi_T{}^*"] \\
 & H^2(Y_T,\boldsymbol{R}).
\end{tikzcd}$$

While $\tau'(\mathring{\sigma}/2\pi\sqrt{-1})$ contains the 2-cocycle (closed 2-form) $-\eta_\sigma$, as is immediately known from the definition of the transgression. Hence we infer that $\tau(\mathring{\sigma}/2\pi\sqrt{-1})$ contains the closed 2-form $-\hat{\eta}_\sigma$ which is naturally induced from $-\eta_\sigma$ on Y_T since $\tau' = \varphi_T{}^* \circ \tau$. This proves the last part of our proposition for $\eta_\sigma = \eta(B_\sigma)$.

Next we want to show that η_σ is of type (1.1): Denote now by ω_α (resp. ω'_α) the left-invariant holomorphic 1-form on G^c (for any root $\mathring{\alpha}$), whose restriction to $\mathfrak{g}^c$ is annihilated by $\mathfrak{t}^c_{\circ}$, $\omega_\alpha(e_\alpha) = 1$ (resp. $\omega'_\alpha(e_\alpha) = \sqrt{-1}$) and $\omega_\alpha(e_\beta) = 0$ (resp. $\omega'_\alpha(e_\beta) = 0$) for $\beta \neq \mathring{\alpha}$. Then the exterior product $\omega_\alpha \wedge \omega_{-\alpha}$ (resp. $\omega'_\alpha \wedge \omega'_{-\alpha}$) for any positive root $\mathring{\alpha}$ is invariant under Ad T^c. Hence the restriction of $\omega_\alpha \wedge \omega_{-\alpha}$ (resp. $\omega'_\alpha \wedge {}'\omega_{-\alpha}$) to G yields the G-invariant (closed) 2-form on $G/T = X_T$, which we will denote by the same symbol $\omega_\alpha \wedge \omega_{-\alpha}$ (resp. $\omega'_\alpha \wedge {}'\omega_{-\alpha}$). We shall show that $\omega_\alpha \wedge \omega_{-\alpha}$ (resp. $\omega'_\alpha \wedge \omega'_{-\alpha}$) is of type (1.1) in the complex manifold X_T. Recall now that the G-invariant complex structure of X_T is induced from that of the open set GU_0 of G^c such that $j_T(X_T) = GU_0/U_0 \subset G^c/U_0 = X_T{}^u$ (cf. § 3, 8). While, the tangent space of $X_T{}^u$ at $x_1{}^u$ (resp. the tangent space of X_T at x_1) is naturally identified with $\mathfrak{p} + \mathfrak{m}$ (resp. with $\mathfrak{p} + \sqrt{-1}\cdot\mathfrak{m}$) (cf. § 1, 2); under these identifications j_T induces the linear mapping j'_T on these tangent spaces such that

$$j'_T \colon x + \sqrt{-1}\cdot y \to x + y \quad (\text{for } x \in \mathfrak{p}, y \in \mathfrak{m}).$$

---

[11] As to the bundles $B_\sigma{}^u$ and $B_\lambda{}^u$, Proposition 4 is already known (cf. Theorem 10. 3 in [4]); hence the proofs shall be omitted for them.

Further we see that the (complex) tangent vectors at x_1^u which annihilate all local anti-holomorphic functions of x_1^u (resp. all local holomorphic functions at x_1^u) are considered, under the above identifications, to span $\mathfrak{n}_0^+$ (resp. $\mathfrak{n}_0^-$). Hence the (complex) tangent vectors at x_1 of X_T which annihilate all local anti-holomorphic functions at x_1 (resp, all local holomorphic functions at x_1) are considered to span $(\mathfrak{n}_0^+ \cap \mathfrak{k}^c) \dotplus \mathfrak{n}^-$ (resp. $(\mathfrak{n}_0^- \cap \mathfrak{k}^c) \dotplus \mathfrak{n}^+$). On the other hand, we have

$$(\omega_\alpha \wedge \omega_{-\alpha})(e_\beta, e_\gamma) = 0 \text{ unless } \mathring{\beta} = -\mathring{\gamma} = \pm\mathring{\alpha}$$

$$(\omega_\alpha \wedge \omega_{-\alpha})(e_\alpha, e_{-\alpha}) = \tfrac{1}{2}.$$

This shows that the left-invariant 2-form $\tilde{\eta}_\sigma$ on G^c is given by

$$\tilde{\eta}_\sigma = \frac{\sqrt{-1}}{2\pi} \sum_{\alpha>0} (\mathring{\sigma}, \mathring{\alpha})\, \omega_\alpha \wedge \omega_{-\alpha},$$

since $\tilde{\eta}_\sigma(e_\alpha, e_{-\alpha}) = \dfrac{\sqrt{-1}}{4\pi}\mathring{\sigma}(h\alpha)$ and $\tilde{\eta}_\sigma(e_\beta, e_\gamma) = 0$ unless $\mathring{\beta} = -\mathring{\gamma} = \pm\mathring{\alpha}$ (cf. the definition of $\tilde{\eta}_\sigma$ in **13**). From this we obtain the formula:

$$\eta_\sigma = \frac{\sqrt{-1}}{2\pi}\{ \sum_{e_\alpha \in k^c} (\mathring{\sigma}, \mathring{\alpha})\, \omega_\alpha \wedge \omega_{-\alpha} - \sum_{e_\beta \in n^+} (\mathring{\sigma}, \mathring{\beta})\, \omega'_\beta \wedge \omega'_{-\beta}\};$$

where we note that ω_α (resp. ω'_β) is the form not defined all over X_T but is defined in a suitable neighbourhood of x_1 as a *holomorphic* 1-form, for which $\omega_{-\alpha}$ (resp. $\omega'_{-\beta}$) represents there the form conjugate to ω_α (resp. to ω'_β). Thus, η_σ is a G-invariant 2-form of type (1.1). While, the form $\hat{\eta}_\sigma$ has the same local expression as η_σ; this proves our proposition for $\hat{\eta}_\sigma = \eta(B_\sigma)$.

Moreover we see that η_σ is the imaginary part of the hermitian form

$$\frac{1}{\pi}\{ \sum_{e_\alpha \in k^c} (\mathring{\sigma}, \mathring{\alpha})\, \omega_\alpha \circ \omega_{-\alpha} - \sum_{e_\beta \in n^+} (\mathring{\sigma}, \mathring{\beta})\, \omega'_\beta \circ \omega'_{-\beta}\},$$

where $\omega_\alpha \circ \omega_{-\alpha}$ (resp. $\omega'_\beta \circ \omega'_{-\beta}$) denotes the G-invariant symmetric differential form on X_T induced from the G^c-invariant form on G^c which is the symmetric product of ω_α and $\omega_{-\alpha}$ (resp. the symmetric product of ω'_β and $\omega'_{-\beta}$). The form $\omega_\alpha \circ \omega_{-\alpha}$ (resp. $\omega'_\beta \circ \omega'_{-\beta}$) coincides, as is $\omega_\alpha \wedge \omega_{-\alpha}$ (resp. $\omega'_\beta \wedge \omega'_{-\beta}$), with the symmetric product of the holomorphic 1-form ω_α and its conjugate $\omega_{-\alpha}$ (resp. ω'_β and $\omega'_{-\beta}$) in a neighbourhood of x_1 in X_T. The explicit expression of η_σ obtained above now yields the

PROPOSITION 5. *The line bundle $F_\sigma = F_{\lambda,\mu}$ is positive in the sense of Kodaira* [14], *if $(\mathring{\sigma}, \mathring{\alpha}) < 0$ for all complementary positive roots $\mathring{\alpha}$ and $(\mathring{\mu}, \mathring{\alpha}_i) > 0$ for all $\mathring{\alpha}_i \in \mathcal{S}$ (while F_σ^u is positive if $(\mathring{\sigma}, \mathring{\alpha}_i) > 0$ for all simple roots $\mathring{\alpha}_i$* [4]).

15. In the case $\sigma = \lambda \in \mathscr{X}_K$, the above arguments go well for the line bundle $E_{\lambda \circ J}$ and its canonical Chern form η_λ (more easily than for $\hat{\eta}_\sigma$) and the detailed proof will be omitted. In fact, for $e_\beta \in \mathfrak{n}^+$, the restriction of $\omega'_\beta \wedge \omega'_{-\beta}$ to G will be readily verified to yield the G-invariant closed 2-form on $G/K = X$, which is also of type (1.1). This form is to be also denoted by the same symbol $\omega'_\beta \wedge \omega'_{-\beta}$; then we can regard it, in a suitable neighbourhood of x_0, as the exterior product of the *holomorphic* 1-form ω'_β and its complex conjugate $\omega'_{-\beta}$. Then we obtain immediately the formula:

$$\eta_\lambda = -\frac{\sqrt{-1}}{2\pi} \sum_{e_\beta \in n^+} (\mathring{\lambda}, \mathring{\beta})\, \omega'_\beta \wedge \omega'_{-\beta}.$$

Further we see that $\hat{\eta}_\lambda$ belongs to the real Chern class of B_λ in a similar way to the case for $\hat{\eta}_\sigma$ and B_σ. Hence

PROPOSITION 6. *The line bundle $E_{\lambda \circ J}$ over Y is positive, if $(\mathring{\lambda}, \mathring{\alpha}_i) < 0$ for all $\mathring{\alpha}_i \notin \mathscr{S}$ (while E_{λ^*} is positive, if $(\mathring{\lambda}, \mathring{\alpha}_i) > 0$ for all $\mathring{\alpha}_i \notin \mathscr{S}$ [4]).*

§ 6. Generalization of Hirzebruch's proportionality principle.

16. Let $\tilde{\eta}$ denote an alternating $2p$-form on $\mathfrak{g}^c$ which represents a $2p$-cocycle in thet sense of the relative cohomology on $\mathfrak{g}^c \bmod \mathfrak{k}^c$ $(1 \leqq p \leqq N)$; we assume further that $\tilde{\eta}$ is *real* valued on $\mathfrak{g}^u$. For such a $\tilde{\eta}$, we denote by η^u (resp. by η) the restriction of $\tilde{\eta}$ to $\mathfrak{g}^u$ (resp. to $\mathfrak{g}$); then both η^u and η are real valued; in fact we see

i) $\eta^u(u_1, \cdots, u_{2p}) = \eta(v_1, \cdots, v_{2p}) = 0$, if there exist elements of $\mathfrak{k}$ among u_i $(1 \leqq i \leqq 2p)$ and among v_i $(1 \leqq i \leqq 2p)$ respectively.

ii) $\eta^u(u_1, \cdots, u_{2p}) = (-1)^p \eta(\sqrt{-1} \cdot u_1, \cdots, \sqrt{-1} \cdot u_{2p})$ for $u_i \in \mathfrak{m}$ $(1 \leqq i \leqq 2p)$.

Hence η^u (resp. η) represents a left invariant closed $2p$-form on $X^u = G^u/K$ (resp. on $X = G/K$). Furthermore η induces naturally the closed $2p$-form $\hat{\eta}$ on $Y = \Gamma \backslash X$. Recall now that η^u is necessarily of type (p, p) (cf. [4])[12]; η is also of type (p, p) as is readily seen (compare with § **5**, **15**). Hence $\hat{\eta}$ is of type (p, p).

Suppose now that $\eta_1^u, \cdots, \eta_s^u$ are G-invariant forms on X^u such that η_i^u is of type (p_i, p_i) $(1 \leqq i \leqq s)$ and $\sum_{i=1}^{s} p_i = N$. We consider the corresponding G-invariant forms $\eta_1, \cdots, \eta_s$ on X and further the closed forms $\hat{\eta}_1, \cdots, \hat{\eta}_s$ over Y. Then the exterior product $\eta_1^u \wedge \cdots \wedge \eta_s^u$ is a G^u-invariant closed form of type (N, N) on X^u, and $\hat{\eta}_1 \wedge \cdots \wedge \hat{\eta}_s$ is a closed form of type (N,N) on Y. Furthermore we take another set of G^u-invariant forms $\zeta_1^u,\cdots,\zeta_t^u$ on X^u with the analogous properties and the corresponding forms $\hat{\zeta}_1,\cdots,\hat{\zeta}_t$ on Y. Then, since $\zeta_1^u\wedge\cdots\wedge\zeta_t^u$ is also a G^u-invariant form of type (N,N) on X^u, there exists a real constant a such that

---

[12] This can be also directly verified.

$$\zeta_1^u\wedge\cdots\wedge\zeta_t^u=a\cdot\eta_1^u\wedge\cdots\wedge\eta_s^u \text{ (on } X^u).$$

By the properties i), ii), we have

$$\zeta_1\wedge\cdots\wedge\zeta_t=a\cdot\eta_1\wedge\cdots\wedge\eta_s \text{ (on } X).$$

Hence

$$\hat{\zeta}_1\wedge\cdots\wedge\hat{\zeta}_t=a\cdot\hat{\eta}_1\wedge\cdots\wedge\hat{\eta}_s \text{ (on } Y).$$

17. Now let $\rho_1,\cdots,\rho_s$ be a finite number of holomorphic representations of K^c. We consider the corresponding bundles $B_{\rho_1}{}^u,\cdots,B_{\rho_s}{}^u$ over X^u, and $B_{\rho_1},\cdots,B_{\rho_s}$ over Y. Let η_i^u (resp. $\hat{\eta}_i$) $(1\leqq i\leqq s)$ denote the p_i-th canonical Chern forms of $B_{\rho_i}{}^u$ (resp. of B_{ρ_i}), assuming that $\sum_{i=1}^{s}p_i=N$. Then we have seen in § 5 that $\hat{\eta}_i$ $(1\leqq i\leqq s)$ are obtained from η_i^u $(1\leqq i\leqq s)$ in the process described above.

By $\mathfrak{z}_{X^u}$ (resp. $\mathfrak{z}_Y$) denotes the generator of the infinite cyclic group $H^{2N}(X^u,\boldsymbol{Z})$ (resp. $H^{2N}(Y,\boldsymbol{Z})$) which is determined by the natural orientation of the complex manifold X^u (resp. Y). Then we can put

$$c_{p_1}(B_{\rho_1}{}^u)\cdots c_{p_s}(B_{\rho_s}{}^u)=c_{X^u}(p_i;\rho_i)\cdot\mathfrak{z}_{X^u},$$
$$c_{p_1}(B_{\rho_1})\cdots c_{p_s}(B_{\rho_s})=c_Y(p_i;\rho_i)\cdot\mathfrak{z}_Y,$$

where $c_{X^u}(p_i;\rho_i)$, $c_Y(p_i;\rho_i)\in\boldsymbol{Z}$. The numbers $c_{X^u}(p_i;\rho_i)$ (resp. $c_Y(p_i;\rho_i)$) are called *Chern numbers*[18] of the bundles $B_{\rho_i}{}^u$ (resp. B_{ρ_i}) $(1\leqq i\leqq s)$). We want now to show the following:

LEMMA 6. *$c_Y(p_i;\rho_i)=\chi(Y)\cdot c_{X^u}(p_i,\rho_i)$, where $\chi(Y)$ denotes the arithmetic genus of Y.*

Proof. It is sufficient to show that there is a rational constant b not depending on the choices of p_i's and ρ_i's such that

$$c_Y(p_i;\rho_i)=b\cdot c_{X^u}(p_i;\rho_i).$$

In fact, if this is the case, we have in particular

$$c_Y(p_i;\mathrm{Ad}_K)=b\cdot c_{X^u}(p_i;\mathrm{Ad}_K)$$

(all ρ_i are Ad_k), while Riemann-Roch-Hirzebruch's theorem implies that the arithmetic genera $\chi(Y)$ and $\chi(X^u)$ are the same linear combinations over $\boldsymbol{Q}$ of these Chern numbers $c_Y(p_i;\mathrm{Ad}_K)$ and of $c_{X^u}(p_i;\mathrm{Ad}_K)$ respectively. Hence we see $\chi(Y)=b\cdot\chi(X^u)=b$, by using $\chi(X^u)=1$ (cf. [4]). Take now another set of holomorphic representations $\phi_1,\cdots,\phi_t$ of K^c, and natural numbers $q_1,\cdots,q_t$ such that $\sum_{j=1}^{t} q_j=N$; we put

---

[18] When all the ρ_i $(1\leqq i\leqq s)$ coincide with $(\mathrm{Ad}_K,\mathfrak{n}^+)$, these are Chern numbers of X^u (resp. of Y) in the usual sense (cf. [9; 10]).

$$c_{X^u}(q_j;\phi_j)=a\cdot c_{X^u}(p_i;\rho_i) \quad (a\in\boldsymbol{Q}).$$

From this, it follows that $\prod_{j=1}^{t}\eta^{q_j}(B_{\phi_j}{}^u)=a\cdot\prod_{i=1}^{s}\eta^{p_i}(B_{\rho_i}{}^u)$ and that $\prod_{j=1}^{t}\eta^{q_j}(B_{\phi_j})$ $=a\prod_{i=j}^{s}\eta^{p_i}(B_{\rho_i})$. Hence we have

$$c_Y(q_j;\phi_j)=a\cdot c_Y(p_i;\rho_i).$$

Thus we see that $c_Y(p_i;\rho_i)/c_{X^u}(p_i;\rho_i)=b$ does not depend on the choices of p_i's and ρ_i's. Our lemma is now proved.

Now, for any $\rho\in\mathrm{Hom}(K^c, GL(m,\boldsymbol{C}))$ we consider the Euler characteristics $\chi(Y,\boldsymbol{E}_{\rho\circ J})$ and $\chi(X^u,\boldsymbol{E}_{\rho^u})$:

$$\chi(Y,\boldsymbol{E}_{\rho\circ J})=\sum_{q=0}^{N}(-1)^q\dim_{\boldsymbol{C}}H^q(Y,\boldsymbol{E}_{\rho\circ J}),$$

$$\chi(X^u,\boldsymbol{E}_{\rho^u})=\sum_{q=0}^{N}(-1)^q\dim_{\boldsymbol{C}}H^q(X^u,\boldsymbol{E}_{\rho^u}).$$

Then, the generalized Riemann-Roch theorem of Hirzebruch [9], combined with Proposition 3, implies that $\chi(Y,\boldsymbol{E}_{\rho\circ J})$ and $\chi(X^u,\boldsymbol{E}_{\rho^u})$ are the same linear combinations of the Chern numbers $c_Y(p_i;\rho_i)$ and $c_{X^u}(p_i;\rho_i)$ respectively, where ρ_i's coincide with ρ or with Ad_K. Therefore, from Lemma 6, we can deduce the following generalization of Hirzebruch's results (see [10; Satz 1 and Satz 4] and also Cor. of Proposition 1).

THEOREM 1. *For any* $\rho\in\mathrm{Hom}(K^c, GL(m,\boldsymbol{C}))$, *it holds*

$$\chi(Y,\boldsymbol{E}_{\rho\circ J})=\chi(Y)\cdot\chi(X^u,\boldsymbol{E}_{\rho^u}).$$

Note that $\chi(X^u,\boldsymbol{E}_{\rho^u})$ is known by the theorem of Bott (cf. § 7, Theorem B$_1$), while $\chi(Y)$ is computable if we know the total volume of Y with respect to a certain invariant metric of X, say the Bergmann's metric. In fact, when X is irreducible, Hirzebruch [11] showed the formula:

$$\chi(Y)=\left(-\tfrac{1}{2\pi}\right)^N\frac{\prod_{\mathring{\alpha}}(\mathring{\delta},\mathring{\alpha})}{(\mathring{\delta}_K,\mathring{\alpha}_k)^N}v(Y),$$

where $\mathring{\alpha}$ runs through all the complementary positive roots, $\mathring{\alpha}_k$ denotes the unique complementary simple root and $v(Y)$ the total volume of Y with respect to the Bergmann's metric of X.

§7. Vanishing theorems for the cohomology groups $H^q(Y, E_{\rho_0, J})$.

18. First we shall resume Bott's main results concerning the induced representations with respect to homogeneous vector bundles [5].

Suppose that $\rho \in \mathrm{Hom}(K^c, GL(m, \boldsymbol{C}))$ is an irreducible representation; for this we consider the homogeneous vector-bundle $E_\rho{}^u$ over X^u and the cohomology groups $H^q(X^u, E_\rho{}^u)$ $(0 \leqq q \leqq N)$. The action of $g \in G^c$ on $E_\rho{}^u$ as bundle automorphism now induces the linear automorphism on the $\boldsymbol{C}$-module $H^q(X^u, E_\rho{}^u)$; which we will write $\rho^{(q)}(g)$. The representation $(\rho^{(q)}, H^q(X^u, E_\rho{}^u))$ of G^c thus obtained is, by definition, the *q-th induced representation of* ρ. Let σ be the highest weight of ρ, then the induced representations $\rho^{(q)}$ $(0 \leqq q \leqq N)$ are determined only by σ:

THEOREM B$_1$. *If there exists a root* $\mathring{\alpha}$ *such that* $(\mathring{\sigma} + \frac{1}{2}\mathring{\delta}, \mathring{\alpha}) = 0$, *then all* $\rho^{(q)}$ $(0 \leqq q \leqq N)$ *are the* 0*-representations* (i.e. $H^*(X^u, E_\rho{}^u) = 0$). *Otherwise, there is one and only one non-zero induced representation* $\rho^{(q)}$ *which is irreducible and its highest weight* $\mathring{\sigma}'$ *is given by*

$$\mathring{\sigma}' + \tfrac{1}{2}\mathring{\delta} = \epsilon(\mathring{\sigma} + \tfrac{1}{2}\mathring{\delta}),$$

where ϵ is the element of Weyl group which is the product of q-reflections with respect to the simple root planes $\mathring{\alpha}_i = 0$ and it is uniquely defined by the condition $(\mathring{\sigma}' + \frac{1}{2}\mathring{\delta}, \mathring{\alpha}_i) > 0$ $(1 \leqq i \leqq l)$.

Alternatively let $\sigma \in \mathrm{Hom}(T^c, \boldsymbol{C}^*) = \mathrm{Hom}(U_1, \boldsymbol{C}^*)$ and consider the homogeneous line bundle $F_\sigma{}^M$ over the manifold $M = K/T = K^c/U_1$ (for the precise definition of $F_\sigma{}^M$, see §4, 9). K^c acts on $F_\sigma{}^M$ as bundle automorphisms and so it induces the representation $\sigma^{(s)}$ of K^c over the s-dimensional cohomology group $H^s(M, \boldsymbol{F}_\sigma{}^M)$ $(1 \leqq s \leqq n = \dim_{\boldsymbol{C}} M)$; $\sigma^{(s)}$ is the *s-th induced representation of* σ. We write now as $\mathring{\sigma} = \mathring{\lambda} + \mathring{\mu}$; $(\mathring{\lambda}, \mathring{\alpha}_i) = 0$ for all $\mathring{\alpha}_i \in \mathscr{S}$ and $(\mathring{\mu}, \mathring{\alpha}_j) = 0$ for all $\mathring{\alpha}_j \notin \mathscr{S}$.

THEOREM B$_2$. *If there exists a root* $\mathring{\alpha}$ *of* $[\mathfrak{k}^c, \mathfrak{k}^c]$ *such that* $(\mathring{\sigma} + \frac{1}{2}\mathring{\delta}, \mathring{\alpha}) = 0$, *or what is the same,* $(\mathring{\mu} + \frac{1}{2}\mathring{\delta}_M, \mathring{\alpha}) = 0$, *then all* $\sigma^{(s)}$ $(1 \leqq s \leqq n)$ *are the* 0*-representations. While, if there is no such a root,* $\sigma^{(s)}$ *are* 0*-representations except for one s; and the remaining representation* $(\sigma^{(s)}, H^s(M, \boldsymbol{F}_\sigma{}^M))$ *is an irreducible one with the highest weight* $\lambda + \mu'$, μ' *being a dominant weight of* $[\mathfrak{k}^c, \mathfrak{k}^c]$ *defined by*

$$\mathring{\mu}' + \tfrac{1}{2}\mathring{\delta}_M = \epsilon'(\mathring{\mu} + \tfrac{1}{2}\mathring{\delta}_M),$$

where ϵ' *is the element of Weyl group of* $[\mathfrak{k}^c, \mathfrak{k}^c]$ *which is the product of* s-*reflections with respect to the simple root planes* $\mathring{\alpha}_i = 0$ $(\mathring{\alpha}_i \in \mathcal{S})$, *and it is uniquely determined by the condition* $(\mathring{\mu}' + \frac{1}{2}\mathring{\delta}_M, \mathring{\alpha}_i) > 0$ *for all* $\mathring{\alpha}_i \in \mathcal{S}$.

The case $s = 0$ in this theorem was first discussed by Borel-Weil [3].

19. Let ρ be an irreducible holomorphic representation of K over $\boldsymbol{C}^m$ with the highest weight $\mathring{\sigma}$; for $\mathring{\sigma}$ we will now define $\lambda \in \mathfrak{X}_K$ and $\mu \in \mathfrak{X}_M$ as in § **4**:

$$\mathring{\sigma} = \mathring{\lambda} + \mathring{\mu}; \; \mathring{\lambda} = \sum_{\alpha_i \notin \mathcal{S}} p_i \mathring{\Lambda}_i, \; \mathring{\mu} = \sum_{\alpha_j \in \mathcal{S}} p_j \mathring{\Lambda}_j.$$

Then our vanishing theorems with respect to $E_{\rho \circ J}$ are formulated as follows:

THEOREM 2. i) *Suppose that*

$$(\check{\lambda} - \mathring{\mu} - \mathring{\delta}_M, \mathring{\alpha}) > 0$$

for all complementary positive roots $\mathring{\alpha}$. *Then*

$$H^q(Y, \boldsymbol{E}_{\rho \circ J}) = 0 \text{ for all } q < N.$$

ii) *Suppose that*

$$(\mathring{\sigma} + \mathring{\delta}, \mathring{\alpha}) < 0$$

for all complementary positive roots $\mathring{\alpha}$. *Then*

$$H^q(Y, \boldsymbol{E}_{\rho \circ J}) = 0 \text{ for all } q > 0.$$

On the other hand, from Bott's theorem (Theorem B_1) we infer easily the following: In the case i), $H^q(X^u, \boldsymbol{E}_{\rho^u}) = 0$ for all $q > 0$ and the 0-th induced representation $\rho^{(0)}$ $(= \rho)$ has the highest weight $\mathring{\sigma}$, while in the case ii), $H^q(X^u, \boldsymbol{E}_{\rho^u}) = 0$ for all $q < N$ and the N-th induced representation $\rho^{(N)}$ has the highest weight $\mathring{\sigma}' - \mathring{\delta}_k$ ($\mathring{\sigma}'$ denotes the highest weight of the contragredient representation of ρ). Thus, combining Theorem 2 with Theorem 1, we can compute the dimensions of the remaining cohomology groups in the respective cases:

COROLLARY. *In the case* i) *we have*

$$\dim_C H^N(Y, \boldsymbol{E}_{\rho \circ J}) = (-1)^N \chi(Y) \deg(\mathring{\sigma}).$$

While, in the case ii) *we have*

$$\dim_C H^0(Y, \boldsymbol{E}_{\rho \circ J}) = (-1)^N \chi(Y) \deg(\mathring{\sigma}' - \mathring{\delta}_k).$$

For the proof of Theorem 2, we need Kodaira's vanishing theorem (cf. [14], and also [9, p. 139]) and the spectral sequence arguments with respect to the projection $\varpi : Y_T \to Y$. We shall prove in the next section the

PROPOSITION 7. *Denoting by* $\varpi^{(s)}(\mathbf{F}_\sigma)$ *the s-dimensional image sheaf*[14] *of the sheaf* $\boldsymbol{F}_\sigma$ *with respect to* ϖ, *it holds that*

$$\varpi^{(s)}(\boldsymbol{F}_\sigma) \cong \boldsymbol{E}_{\sigma^{(s)}\circ J}$$

as (*analytic*) *sheaves on* Y, *where* $\sigma^{(s)}$ *is the holomorphic representation of* K^c *over the representation module* $H^s(M, \boldsymbol{F}_\sigma{}^M)$ *induced by* $\mathring{\sigma}$ *in the sense of Bott* (cf. Theorem B_2).

We know that $\sigma^{(s)}$ $(0 \leqq s \leqq n = \dim_C M)$ are zero representations except for at most one value of s (cf. Theorem B_2). It follows therefore from a theorem of Leray [5, p. 237] that the canonical spectral sequence $\{E_k\}$ is trivial; here $\{E_k\}$ denotes the spectral sequence whose final term E_∞ is associated to H^* $(Y_T, \boldsymbol{F}_\sigma)$ and whose E_2-term is given by $E_2{}^{r,s} = H^r(Y, \varpi^{(s)}(\boldsymbol{F}_\sigma))$. Hence, an immediate consequence of Proposition 7 is the following:

PROPOSITION 8. i) *Suppose now that* $\sigma^{(s)}$ *is the only non-zero induced representation of* σ *among those for* $0 \leqq s \leqq n$. *Then we have*

$$H^q(Y, \boldsymbol{E}_{\rho\circ J}) \cong H^{q+s}(Y_T, \boldsymbol{F}_\sigma), \quad \textit{for} \quad 0 \leqq q \leqq N,$$

where $\rho = \sigma^{(s)} = \lambda \otimes \mu^{(s)}$ *for* $\sigma = (\lambda, \mu)$.

ii) *In case all* $\sigma^{(s)}$ $(0 \leqq s \leqq n)$ *are zero representations, we have*

$$H^q(Y_T, \boldsymbol{F}_\sigma) = 0, \quad \textit{for} \quad 0 \leqq q \leqq N + n.$$

We shall now proceed to the proof of Theorem 2 by using Proposition 8. First we show ii): We set $\rho' = \lambda^{-1} \otimes \rho$, then μ is the highest character of ρ', and $\mu \in \mathscr{X}_M$. The theorem of Borel-Weil-Bott (cf. Theorem B_2) implies that $\rho' = \mu^{(0)}$ (i.e. $\rho = \sigma^{(0)}$); hence from i), Proposition 8, we infer that

$$H^q(Y, \boldsymbol{E}_{\rho\circ J}) \cong H^q(Y_T, \boldsymbol{F}_\sigma), \quad 0 \leqq q \leqq N.$$

On the other hand, $K_T{}^* \otimes F_\sigma = F_{\sigma\cdot\delta}$ (cf. Corollary of Lemma 4); where we see that $(\mathring{\sigma} + \mathring{\delta}, \mathring{\alpha}_i) = (\mathring{\mu} + \mathring{\delta}_M, \mathring{\alpha}_i) > 0$ for all $\alpha_i \in \mathscr{S}$ since $\mathring{\mu}$ is a dominant weight in $[\mathfrak{k}^c, \mathfrak{k}^c]$. While, by assumption, $(\breve{\sigma} + \mathring{\delta}, \mathring{\alpha}) < 0$ for all complementary root $\mathring{\alpha}$. Therefore, by Proposition 5 we see that $K_T{}^* \otimes F_\sigma$ is a positive line bundle; hence, by Kodaira's vanishing theorem [15], we have $H^q(Y_T, \boldsymbol{F}_\sigma) = 0$ for all $q > 0$. This proves ii).

Next, to prove i), for $\rho = \lambda \otimes \rho'$ we put

$$\mathring{\nu} = -\mathring{\mu} - \mathring{\delta}_M \quad (\nu \in \mathscr{X}_M).$$

Then $(\mathring{\lambda} + \mathring{\nu}, \mathring{\alpha}_i) = (\mathring{\nu}, \mathring{\alpha}_i) < 0$ for all $\mathring{\alpha}_i \in \mathscr{S}$, while by assumption $(\mathring{\lambda} + \mathring{\nu}, \mathring{\alpha}) = \mathring{\lambda} - \mathring{\mu} - \mathring{\delta}_M, \mathring{\alpha}) > 0$ for all complementary roots $\mathring{\alpha}$. Hence, Proposition 5 implies that $F_{\lambda,\nu}$ is a negative line bundle, and so we infer from Kodaira's theorem that

---

[14] The definition of the s-dimensional image sheaf will be recalled in the beginning of §8.

$$H^q(Y_T, \boldsymbol{F}_{\lambda,\nu}) = 0 \text{ for all } q < n + N.$$

On the other hand, Bott's theorem (cf. Theorem B_2) now implies that

$$(\rho', \boldsymbol{C}^m) \sim (\nu^{(n)}, H^n(M, \boldsymbol{F}_\nu{}^M))\,;$$

so that it follows from Proposition 8 that

$$H^q(Y, \boldsymbol{E}_{\rho\circ J}) \cong H^{q+n}(Y_T, \boldsymbol{F}_{\lambda,\nu}), \quad 0 \leqq q \leqq N.$$

Hence, combining with the above equality, we obtain the desired result. The proof of Theorem 2 is now complete.

20. In case σ satisfies neither the condition in i) nor the one in ii), Theorem 2 tells us nothing about the cohomology groups $H^q(Y, \boldsymbol{E}_{\rho\circ J})$. However we can establish the following result by using Bott's vanishing theorem [5, p. 231; Cor.], in place of Kodaira's one:

THEOREM 3. *If there exists at least one complementary positive root* $\mathring{\alpha}$ *such that* $(\mathring{\sigma}, \mathring{\alpha}) > 0$, *then it holds that*

$$H^0(Y, \boldsymbol{E}_{\rho\circ J}) = 0$$

Proof. Proposition 8 implies, as in ii) of Theorem 2, that

$$H^0(Y, \boldsymbol{E}_{\rho\circ J}) \cong H^0(Y_T, \boldsymbol{F}_\sigma).$$

The condition of our theorem and the proof of Proposition 4 imply that the canonical Chern form of F_σ has one eigen value, corresponding to $\mathring{\alpha}$, which is negative throughout Y_T. Hence, from Bott's theorem quoted above, it follows our theorem.

Remark 1. Our vanishing theorems (i), ii) in Theorem 2) are stated for the m-dimensional vector bundle $E_{\rho\circ J}$ $(m \geqq 1)$; however, in the case of the line bundle $E_{\lambda\circ J}$, the conditions on $\mathring{\lambda}$ will be revised from those in the above theorems. In fact, as for the cohomology groups $H^q(Y, \boldsymbol{E}_{\lambda\circ J})$, we can immediately apply Kodaira's vanishing theorem by using Proposition 6 and Cor. of Proposition 1. Thus we get a sharpened form of Hirzebruch's results [10, § 3]:

i) *If* $(\lambda, \alpha_i) > 0$ *for all* $\alpha_i \notin \mathcal{S}$, *then* $H^q(Y, \boldsymbol{E}_{\lambda\circ J}) = 0$ $(q < N)$, *and*

ii) *If* $(\lambda + \delta_K, \alpha_i) < 0$ *for all* $\alpha_i \notin \mathcal{S}$, *then* $H^q(Y, \boldsymbol{E}_{\lambda\circ J}) = 0$ $(q > 0)$.

We recall that Hirzebruch has been concerned in [10] with the case $\mathring{\lambda} = r \cdot \mathring{\delta}_K$ $(r \in \boldsymbol{Z})$.

Remark 2. In the above arguments, we have assumed that ρ is irreducible.

While, in the reducible case, we let $\rho = \rho_1 \oplus \rho_2 \oplus \cdots \oplus \rho_k$ denote the decomposition of ρ into the irreducible components. Then we infer easily from the definition of $E_{\rho \circ J}$ that $E_{\rho \circ J} = E_{\rho_1 \circ J} \oplus E_{\rho_2 \circ J} \oplus \cdots \oplus E_{\rho_k \circ J}$ ($\oplus$ denotes the Whitney sum). Hence we have

$$H^q(Y, \boldsymbol{E}_{\rho \circ J}) = \sum_{i=1}^{k} H^q(Y, \boldsymbol{E}_{\rho_i \circ J}), \quad 0 \leqq q \leqq N.$$

Remark 3. We add here that the cohomology groups $H^q(Y, \boldsymbol{\Theta})$ have been already discussed by Calabi-Vesentini [6] by using the theory of harmonic forms; however, these cannot be computed by our method.

§ 8. The proof of Proposition 7.

21. Throughout this section, it will be meant by N an (arbitrary) open set of Y. We set $\sigma^{(s)} = \rho$, $\sigma = (\lambda, \mu)$.

The sheaf $\varpi^{(s)}(\boldsymbol{F}_{\lambda,\mu})$ is by definition the one associated to the presheaf $\{\varpi^{(s)}(\boldsymbol{F}_{\lambda,\mu})_N ; N \subset Y\}$, which is given by

$$\varpi^{(s)}(\boldsymbol{F}_{\lambda,\mu})_N = H^s(\varpi^{-1}(N), F_{\lambda,\mu} \mid \varpi^{-1}(N)).$$

While, the sheaf $\boldsymbol{E}_{\rho \circ J}$ is the one associated to the pre-sheaf $\{(\boldsymbol{E}_{\rho \circ J})_N ; N \subset Y\}$, where

$$(\boldsymbol{E}_{\rho \circ J})_N = H^0(N, \boldsymbol{E}_{\rho \circ J} \mid N).$$

In order to prove that $\varpi^{(s)}(\boldsymbol{F}_{\lambda,\mu}) \cong \boldsymbol{E}_{\rho \circ J}$ as (analytic) sheaves, it suffices to show the existence of a pre-sheaf homomorphism $\{h_N ; N \subset Y\}$ of

$$\{\varpi^{(s)}(\boldsymbol{F}_{\lambda,\mu})_N ; N \subset Y\}$$

into $\{(\boldsymbol{E}_{\rho \circ J})_N ; N \subset Y\}$ which induces an isomorphism on each fibre of the limit sheaves: i.e., h_N is a homomorphism of $\varpi^{(s)}(\boldsymbol{F}_{\lambda,\mu})_N$ into $(\boldsymbol{E}_{\rho \circ J})_N$ such that i) h_N's commute with the restriction maps, and ii) for any point $x \in Y$ and any N containing x there exists an open set $N' \subset N$ $(x \in N')$ on which $h_{N'}$ is bijective.

Recall now that the $\boldsymbol{C}$-module $H^0(N, \boldsymbol{E}_{\rho \circ J} \mid N)$ may be identified with that of all holomorphic mappings s of $\varphi^{-1}(N)$ into $H^s(M, \boldsymbol{F}_\mu{}^M)$ such that

$$s(\gamma \cdot x) = \rho(J(\gamma, x)) \cdot s(x)$$

for $\gamma \in \Gamma$ and $x \in \varphi^{-1}(N)$, since the standard fibre of the vector bundle $E_{\rho \circ J}$ is given by $H^s(M, F_\mu{}^M)$.

Before constructing $\{h_N ; N\}$, we have first to define, for every $x \in \varphi^{-1}(N)$, the bundle mapping $\hat{x}$ of $F_\mu{}^M$ into $F_{\lambda,\mu} \mid \varpi^{-1}(N)$; the latter being identified, as in Lemma 4, with $\Gamma \backslash \varphi^{-1}(N) \times F_\mu{}^M$). In fact we will set:

$$\hat{x}(f) = \varphi_T(x,f) \in F_{\lambda,\mu} \mid \varpi^{-1}(N),$$

for $f \in F_\mu^M$. Then the mapping $\hat{x}$ induces on the base spaces the injective mapping:

$$M \ni z \to \varphi_T(x,z) \in \varpi^{-1}(N).$$

If we set $f = [k,\xi]_\mu$: $k \in K^c$, $\xi \in \boldsymbol{C}$ (cf. § 3), our mapping $\hat{x}$ satisfies then the following relations:

$$(\widehat{\gamma x})(f) = \hat{x}([J(\gamma,x)^{-1}k, \lambda(J(\gamma,x)^{-1})\xi]_\mu),$$

for $\gamma \in \Gamma$, as is seen from Lemma 4. Denote by $\hat{x}^*$ the homomorphism of $H^s(\varpi^{-1}(N), \boldsymbol{F}_{\lambda,\mu} \mid \varpi^{-1}(N))$ into $H^s(M, \boldsymbol{F}_\mu^M)$ which is induced naturally by $\hat{x}$:

$$\hat{x}^*: H^s(\varpi^{-1}(N), \boldsymbol{F}_{\lambda,\mu} \mid \varpi^{-1}(N)) \to H^s(M, \boldsymbol{F}_\mu^M).$$

Now we can put for every $\eta \in H^s(\varpi^{-1}(N), \boldsymbol{F}_{\lambda,\mu} \mid \varpi^{-1}(N))$ as

$$h_N(\eta) = s\,;\, s(x) = \hat{x}^*(\eta), \quad \text{for all } x \in \varphi^{-1}(N).$$

Then, by the above relations, the mapping s of $\varphi^{-1}(N)$ into $H^s(M, \boldsymbol{F}_\mu^M)$ satisfies the relations:

$$s(\gamma\cdot x) = (\widehat{\gamma\cdot x})^*(\eta) = (\lambda \otimes \mu^{(s)})(J(\gamma,x))s(x)$$

for $x \in \varphi^{-1}(N)$, $\gamma \in \Gamma$ and $\eta \in H^s(\varpi^{-1}(N), \boldsymbol{F}_{\lambda,\mu} \mid \varpi^{-1}(N))$. Therefore, for proving that $s \in H^0(N, \boldsymbol{E}_{\rho\circ J} \mid N)$, we have only to verify that the mapping s is *holomorphic*.

Let τ_1 now denote the projection of $X_T = X \times M$ onto M, then we infer that $\varphi_T^* F_{\lambda,\mu} \, (= j_T^* F_{\sigma^u}) \cong \tau_1^* F_\mu^M$ (cf. the proof of Lemma 4); we can thus consider the diagram

$$\begin{tikzcd}
H^s(\varpi^{-1}(N), \boldsymbol{F}_{\lambda,\mu} \mid \varpi^{-1}(N)) \arrow[r, "\hat{x}^*"] \arrow[d, "\varphi_T^*"'] & H^s(M, \boldsymbol{F}_\mu^M) \\
H^s(\varphi_T^{-1}\varpi^{-1}(N), \ \tau_1^* \boldsymbol{F}_\mu^M \mid \varphi_T^{-1}\varpi^{-1}(N)), \arrow[ur, "\tilde{x}^*"'] &
\end{tikzcd}$$

where φ_T^* denotes the natural homomorphism induced by the bundle mapping φ_T: $\varphi_T^* F_{\lambda,\mu} = \tau_1^* F_\mu^M \to F_{\lambda,\mu}$, and $\tilde{x}^*$ is the one induced by the bundle mapping $\tilde{x}$: $F_\mu^M \to \tau_1^* F_\mu^M = \varphi^{-1}(N) \times F_\mu^M$ such that $\tilde{x}(f) = (x,f)$ for $f \in F_\mu^M$. It clearly holds that $\hat{x}^* = \tilde{x}^* \circ \varphi_T^*$. Therefore, in order to show that $x \to s(x) = \hat{x}^*(\eta)$ is a holomorphic mapping, it suffices to show that $x \to \tilde{x}^*(\zeta)$ is a holomorphic one of $\varphi^{-1}(N)$ into $H^s(M, \boldsymbol{F}_\mu^M)$ for any fixed $\zeta \in H^s(\varphi_T^{-1}\cdot \varpi^{-1}(N)), \tau_1^* \cdot \boldsymbol{F}_\mu^M \mid \varphi_T^{-1} \cdot \varpi^{-1}(N))$. Now take a Stein open set

N_0 for a given x such that $x \in \varphi^{-1}(N_0)$, $N_0 \subset N$. Then, since $\varphi_T^{-1}\cdot \varpi^{-1}(N_0) = \pi^{-1}\cdot \varphi^{-1}(N_0) = \varphi^{-1}(N_0)\times M$, the mapping $\tilde{x}^* = \tilde{x}^*(N_0)$ will map

$$H^s(\varphi^{-1}(N_0)\times M, \tau_1^{*}\ \boldsymbol{F}_\mu^M \mid (\varphi^{-1}(N_0)\times M))$$

into $H^s(M, \boldsymbol{F}_\mu^M)$. On the other hand, the Künneth relation implies that

$$H^s(\varphi^{-1}(N_0)\times M, \tau_1{}^*\boldsymbol{F}_\mu^M \mid (\varphi^{-1}(N_0)\times M)) \cong H^0(\varphi^{-1}(N_0), \mathfrak{O}) \otimes H^s(M, \boldsymbol{F}_\mu^M),$$

since $\varphi^{-1}(N_0)$ is also a Stein manifold; *via* this isomorphism $\tilde{x}^*(N_0)$ is represented such that $\tilde{x}^*(N_0)(\alpha\otimes\xi) = \alpha((x)\otimes\xi$, where $\xi \in H^s(M, \boldsymbol{F}_\mu^M)$ and $\alpha \in H^0(\phi^{-1}(N_0), \mathfrak{O})$ is regarded as a holomorphic function on $\varphi^{-1}(N_0)$. Hence $x \to \tilde{x}^*(N_0)(\alpha\otimes\xi)$ is a holomorphic mapping of $\varphi^{-1}(N_0)$ into $H^s(M, \boldsymbol{F}_\mu^M)$. That $s \in H^0(N, \boldsymbol{E}_{\rho\circ J} \mid N)$ is thus proved.

Next, we have to verify the relations:

$$r_{N'}{}^N \circ h_N = h_{N'} \circ r_{N'}{}^N$$

for two arbitrary open sets N and N' $(N \supset N')$, where $r_{N'}{}^N$ denotes the restriction mapping; in other words, it should be checked that the following diagram is commutative:

$$\begin{tikzcd}
\varpi^{(s)}(\boldsymbol{F}_{\lambda,\mu})_N \arrow[r, "h_N"] \arrow[d, "r_{N'}{}^N"] & (E_{\rho\circ J})_N \arrow[d, "r_{N'}{}^N"] \\
\varpi^{(s)}(\boldsymbol{F}_{\lambda,\mu})_{N'} \arrow[r, "h_{N'}"'] & (E_{\rho\circ J})_{N'}
\end{tikzcd}$$

In fact, if we set $s = h_N(\eta)$, $r_{N'}{}^N(s) = s'$ and $r_{N'}{}^N(\eta) = \eta'$ for any $\eta \in \varpi^{(s)}(\boldsymbol{F}_{\lambda,\mu})_N$, then by definition $s(x) = \hat{x}^*(\eta)$ for any $x \in \varphi^{-1}(N)$. It will be seen moreover that s' is the restriction, as a mapping, of s to $\varphi^{-1}(N')$, and that $\hat{x}^*(\eta') = \hat{x}^*(\eta)$ for $x \in \varphi^{-1}(N')$; hence it holds that $s'(x) = \hat{x}^*(\eta')$ for $x \in \varphi^{-1}(N')$. This implies that $s' = h_{N'}(\eta)$.

22. Finally, the pre-sheaf homomorphism $\{h_N; N \subset Y\}$ defined above must be proved to induce in reality an isomorphism in the limit sheaves; namely we must show that, for a given point of Y and its open neighborhood, there is an open neighborhood N contained in it such that h_N yields a bijection of $\varpi^{(s)}(\boldsymbol{F}_{\lambda,\mu})_N$ onto $(\boldsymbol{E}_{\rho\circ J})_N$. Now, we take, as such a neighborhood N, a contractible Stein open set; for which we shall prove that h_N is bijective. In fact, in this case, since the two bundles $\varpi^{-1}(N) \to N$ and $\varphi^{-1}(N) \to N$ are trivial, we can identify such that

$$\varpi^{(s)}(\boldsymbol{F}_{\lambda,\mu})_N = H^s(N\times M, \boldsymbol{F}_{\lambda,\mu} \mid (N\times M)),$$

where $F_{\lambda,\mu} \mid (N\times M) = \tau_1^*\cdot F_\mu^M$ since $E_{\lambda\circ J}\mid N$ is trivial (τ_1 is the projection of $N\times M$ onto M). Hence it holds that

$$\varpi^{(s)}(\boldsymbol{E}_{\lambda,\mu})_N = H^s(N\times M, \tau_1^*\cdot \boldsymbol{F}_\mu^M).$$

On the other hand, $E_{\rho\circ J}\mid N$ being trivial, we may identify

$$(\boldsymbol{E}_{\rho\circ J})_N = \mathfrak{A}_N\otimes H^s(M,\boldsymbol{F}_\mu^M),$$

where $\mathfrak{A}_N$ denotes the $\boldsymbol{C}$-module of all holomorphic functions on N. This identification will be realized as follows: First we know by Proposition 1 that

$$E_{\rho\circ J}\mid N = \Gamma\backslash(\varphi^{-1}(N)\times H^s(M,\boldsymbol{F}_\mu^M)),$$

where we may consider as $\varphi^{-1}(N) = \Gamma\times N$ since $\varphi^{-1}(N)\to N$ is trivial; we write thus $\varphi^{-1}(N)\ni x = (\gamma,y) = \gamma\cdot y$ ($\gamma\in\Gamma$, $y\in N$). Then, for $x=(\gamma,y)\in\varphi^{-1}(N)$ and $\zeta\in H^s(M,\boldsymbol{F}_\mu^M)$, we have the bundle isomorphism:

$$\varphi(x,\zeta)\to(y,\rho(J(\gamma,y))^{-1}\zeta)$$

of $E_{\rho\circ J}\mid N$ onto $N\times H^s(M,\boldsymbol{F}_\mu^M)$; in fact, under this correspondence, $\varphi(\gamma' x;(J(\gamma',x))\zeta)$ is mapped to

$$(y,\rho(J(\gamma'\gamma,y))^{-1}\cdot\rho(J(\gamma',x)\zeta) = (y,\rho(J(\gamma,y))^{-1}\zeta)$$

Recalling now that $(\boldsymbol{E}_{\rho\circ J})_N$ consists of all holomorphic mapping s of $\varphi^{-1}(N)$, into $H^s(M,\boldsymbol{F}_\mu^M)$ such that $s(\gamma x) = \rho(J(\gamma,x))s(x)$ for $x\in\varphi^{-1}(N)$, $\gamma\in\Gamma$, the restriction $s\mid N$ of s to N may be regarded as a holomorphic mapping of N into $H^s(M,\boldsymbol{F}_\mu^M)$:

$$s\mid N\in\mathfrak{A}_N\otimes H^s(M,\boldsymbol{F}_\mu^M).$$

In fact, for any $y\in N$, $\rho(J(\gamma,y))^{-1}\cdot s(\gamma\cdot y) = s(y)(=\zeta)$ and conversely to $\varphi(y,\zeta)\in E_{\rho\circ J}\mid N$ corresponds (y,ζ); hence $y\to\zeta$ gives a holomorphic section of $E_{\rho\circ J}\mid N$.

Now let $\tau_1\colon N\times M\to M$ and $\tau_2\colon N\times M\to N$ be the natural projections, then they induce respectively the natural homomorphisms τ_1^* and τ_2^* such that

$$\tau_1^*\colon H^s(M,\boldsymbol{F}_\mu^M)\to H^s(N\times M,\tau_1^*\boldsymbol{F}_\mu^M)$$

$$\tau_2^*\colon\mathfrak{A}_N\to\mathfrak{A}_{N\times M}=\mathfrak{A}_N$$

($\mathfrak{A}_{N\times M}$ is the module of all holomorphic functions on $N\times M$). From them, we get the homomorphism

$$\tau_2^*\otimes\tau_1^*\colon\mathfrak{A}_N\otimes H^s(M,\boldsymbol{F}_\mu^M)\to H^s(N\times M\cdot\tau_1^*\boldsymbol{F}_\mu^M):$$

while we see that this is a special case of the so-called Künneth relation, so that $\tau_2^*\otimes\tau_1^*$ is bijective (cf. [5], p. 237). Therefore, what is to prove is that

$$\begin{aligned} h_N\circ(\tau_2^*\otimes\tau_1^*) = {} & \text{the identity mapping of} \\ & \mathfrak{A}_N\otimes H^s(M,F_\mu^M) \text{ onto itself.}\end{aligned}$$

To show this, let $\alpha\otimes\zeta\in\mathfrak{A}_N\otimes H^s(M,F_\mu^M)$ ($\alpha\in\mathfrak{A}_N$ and $\zeta\in H^s(M,F_\mu^M)$), and put

$$\begin{gathered}(\tau_2^*\otimes\tau_1^*)(\alpha\otimes\zeta)=\eta\in H^s(N\times M,\tau_1^*F_\mu^M)\\ \{h_N\circ(\tau_2^*\otimes\tau_1^*)\}(\alpha\otimes\zeta)=h_N(\eta)=s.\end{gathered}$$

We will thereby show that $s=\alpha\otimes\zeta$. Recall now that the bundle diagram in our case is as follows:

$$\begin{tikzcd} N\times M & \Gamma\times N\times M \arrow[l, "\varphi_T"'] \arrow[d, "\pi"] \\ M & \Gamma\times N \arrow[l, "\varphi"'] \arrow[from=1-1, to=2-1, "\varpi"'] \end{tikzcd}$$

where Γ acts on $\Gamma\times N$ such that $\gamma\circ(\gamma',y)=(\gamma\gamma',y)$, and does on $\Gamma\times N\times M$ such that

$$\gamma\circ(\gamma',y,z)=(\gamma\gamma',y,J(\gamma,y)z),$$

for $\gamma,\gamma'\in\Gamma$, $y\in N$ and $z\in M$. Thus we see

$$\begin{aligned}s(y)&=\hat{y}^*(\eta)=\hat{y}^*(\tau_2^*\otimes\tau_1^*)(\alpha\otimes\zeta)\\ &=\hat{y}^*(\alpha\otimes\tau_1^*(\zeta))=\alpha(y)(\hat{y}^*\cdot\tau_1^*)(\zeta).\end{aligned}$$

On the other hand, for any $f\in F_\mu^M$ we see that

$$(\tau_1\circ\hat{y})(f)=\tau_1([y,f]_\mu)=f;$$

Hence, $\tau_1\circ\hat{y}$ is the identity. This implies that $(\tau_1\circ y)^*\zeta=\zeta$. Therefore $s(y)=\alpha(y)\zeta=(\alpha\otimes\zeta)(y)$. The proof of Projosition 7 is now complete.

§ 9. The case of Siegel's symplective space.

23. In this section, as an illustration of our main results, we will deal with the case where X is the irreducible symmetric bounded domain of type III_n (the so-called Siegel's symplectic space of degree n). Namely X is the space of all complex symmetric matrices x of degree n such that $x\bar{x}-1_n$ (1_n denotes the unit matrix of degree n) is negative definite (cf. [7]). In this case we can put:

$$G^{c}=S_p(n,\boldsymbol{C}),\qquad G^{u}=LS_p(n)L^{-1},\qquad G=LS_p(n,\boldsymbol{R})L^{-1},$$
$$K=L(S_p(n)\cap S_p(n,\boldsymbol{R}))L^{-1}$$

where $L\in GL(2n,\boldsymbol{C})$ denotes the matrix

$$L=\begin{pmatrix}1_n & -(-1)^{\frac{1}{2}}\cdot 1_n\\ 1_n & (-1)^{\frac{1}{2}}\cdot 1_n\end{pmatrix}.$$

Then K is the maximal compact subgroup of G which consists of matrices

$$\begin{pmatrix}k & 0\\ 0 & {}^{t}k^{-1}\end{pmatrix};\quad w\in U(n),$$

whence K^{c} consists of matrices of the same form such that $k\in GL(n,\boldsymbol{C})$. K^{c} is *canonically* isomorphic to $GL(n,\boldsymbol{C})$ under the correspondence $\begin{pmatrix}k & 0\\ 0 & {}^{t}k^{-1}\end{pmatrix}\to k$. Now we can take the maximal torus T of K such that

$$T=\left\{\begin{pmatrix}a & 0\\ 0 & a^{-1}\end{pmatrix}\,\middle|\, a=\begin{pmatrix}t_1 & & \\ & \ddots & \\ & & t_n\end{pmatrix};\ t_i=e^{\sqrt{-1}\cdot h_i}\ (h_i\in\boldsymbol{R})\right\};$$

then

$$T^{c}=\left\{\begin{pmatrix}a & 0\\ 0 & a^{-1}\end{pmatrix}\,\middle|\, a=\begin{pmatrix}t_1 & & \\ & \ddots & \\ & & t_n\end{pmatrix};\ t_i\in\boldsymbol{C}^{*}\right\},$$

$$\mathfrak{t}^{c}=\left\{\begin{pmatrix}a & 0\\ 0 & -a\end{pmatrix}\,\middle|\, a=\begin{pmatrix}h_1 & & \\ & \ddots & \\ & & h_m\end{pmatrix};\ h_i\in\boldsymbol{C}\right\}.$$

We set here $\mathring{\lambda}_i(h)=h_i\ (1\leqq i\leqq n)$, $\mathring{\alpha}_i=\lambda_i-\lambda_{i+1}\ (1\leqq i\leqq n-1)$, $\mathring{\alpha}_n=2\mathring{\lambda}_n$.[15] Then $\{\mathring{\alpha}_1,\mathring{\alpha}_2,\cdots,\mathring{\alpha}_n\}$ constitutes a simple root system of $\mathfrak{g}=\mathfrak{sp}(n,\boldsymbol{C})$, and $\mathcal{S}=\{\mathring{\alpha}_1,\mathring{\alpha}_2,\cdots,\mathring{\alpha}_{n-1}\}$ a simple root system of $[\mathfrak{k}^{c},\mathfrak{k}^{c}]=\mathfrak{sl}(n,\boldsymbol{C})$. Moreover we know that the fundamental dominant weights system corresponding to $\{\mathring{\alpha}_1,\mathring{\alpha}_2,\cdots,\mathring{\alpha}_n\}$ is given by

$$\mathring{\Lambda}_i=\mathring{\lambda}_1+\mathring{\lambda}_2+\cdots+\mathring{\lambda}_i\quad(1\leqq i\leqq n),$$

and that

$$\mathring{\delta}=2(\mathring{\Lambda}_1+\cdots+\mathring{\Lambda}_n),\quad \mathring{\delta}_M=2(\mathring{\Lambda}_1+\cdots+\mathring{\Lambda}_{n-1}),$$
$$\mathring{\delta}_K=(n+1)\mathring{\Lambda}_n.$$

For the linear order given by the above simple roots system, we infer readily that

---

[15] For the root system and the root vectors of $\mathfrak{sp}(n,C)$, we refer to [13].

$$\mathfrak{n}^+ = \left\{ \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} \mid {}^t b = b \right\}, \quad \mathfrak{n}^- = \left\{ \begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix} \mid {}^t c = c \right\},$$

$$N^+ = \left\{ \begin{pmatrix} 1_n & b \\ 0 & 1_n \end{pmatrix} \mid {}^t b = b \right\}, \quad N^- = \left\{ \begin{pmatrix} 1_n & 0 \\ c & 1_n \end{pmatrix} \mid {}^t c = c \right\}.$$

On the other hand we see that G consists of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that

$$a = \bar{d}, \qquad b = \bar{c}, \qquad {}^t a c = {}^t c a,$$
$$ {}^t \bar{a} a - {}^t c \bar{c} = 1_n.$$

G acts on X transitively by the following rule: for $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$ and $x \in X$.

$$g \cdot x = (ax + b)(x + d)^{-1}.$$

Then, the basic point $x_0 \in X$ is the zero matrix of degree n (K is the isotropy subgroup of G at x_0). We see without difficulties that $G \subset N^+ K^c N^-$, and that

$$j_1(x) = \begin{pmatrix} 1_n & x \\ 0 & 1_n \end{pmatrix}, \quad x \in X;$$

moreover we have (cf. Lemma 1 in § 1)

$$g \cdot j_1(x) = j_1(gx) \begin{pmatrix} {}^t(cx + d)^{-1} & 0 \\ 0 & cx + d \end{pmatrix} \cdot \begin{pmatrix} 1_n & 0 \\ (cx + d)^{-1} \cdot c & 1_n \end{pmatrix}.$$

Thus we can put

$$J(g, x) = {}^t(cx + d)^{-1},$$

after identifying K^c with $GL(n, \boldsymbol{C})$ through the canonical isomorphism considered above. It follows then that the Jacobian of g at x is given by

$$(\delta_K \circ J)(g, x) = \det(cx + d)^{-n-1}.$$

We consider now the irreducible (holomorphic) representation ρ of $GL(n, \boldsymbol{C})$ over $\boldsymbol{C}^m$ with the highest weight $\mathring{\sigma}$; then we may write as

$$\mathring{\sigma} = \mathring{\lambda} + \mathring{\mu}; \ \mathring{\lambda} = p\mathring{\Lambda}_n, \quad \mathring{\mu} = \sum_{i=1}^{n-1} p_i \mathring{\Lambda}_i,$$

where p and p_i $(1 \leqq i \leqq n-1)$ are integers such that $p_i \geqq 0$ $(1 \leqq i \leqq n-1)$ and p may be arbitrary.

Under these circumstances, Corollary of Theorem 2, Theorem 3 and Remark 1 in § 7 now imply:

THEOREM 4. i) *If* $p - \sum_{i=1}^{n-1} p_i > 2n - 2$, *or if* $\sigma = \lambda$ (i. e. $m = 1$) *and* $p > 0$, *then*

$$H^q(Y, \boldsymbol{E}_{\rho\circ J}) = 0 \textit{ for all } q < N,$$

$$\dim_{\boldsymbol{C}} H^N(Y, \boldsymbol{E}_{\rho\circ J}) = (4\pi)^{-N}2^{-n}\prod_{k=1}^{n}\frac{(2k)!}{k!}\cdot v_S(Y)\deg(\mathring{\sigma}),$$

where $v_S(Y)$ *denotes the total volume of* Y *with respect to the Siegel's volume element* (cf. [11]), *and* $\deg(\mathring{\sigma})$ *denotes the degree of the irreducible representation of* $S_p(n, \boldsymbol{C})$ *whose highest weight is* $\mathring{\sigma}$

ii) *If* $p + \sum_{i=1}^{n-1} p_i < -2n$, *or if* $\sigma = \lambda$ *and* $p < -n-1$, *then*

$$H^q(Y, \boldsymbol{E}_{\rho\circ J}) = 0 \textit{ for all } q > 0,$$

$$\dim_{\boldsymbol{C}} H^0(Y, \boldsymbol{E}_{\rho\circ J}) = (4\pi)^{-N}2^{-n}\prod_{k=1}^{n}\frac{k!}{(2k)!}v_S(Y)\deg(\mathring{\sigma}' + \mathring{\delta}_K),$$

where $\mathring{\sigma}$ *denotes the highest weight of the contragredient representation of* ρ.

iii) *If* $p + \sum_{i=1}^{n-1} p_i > 0$ (*in particular, if* $p \geqq 0$ *and* ρ *is non-trivial*), *then*

$$H^0(Y, \boldsymbol{E}_{\rho\circ J}) = 0,$$

§ 10. Final remarks.

24. Throughout the present paper, our methods have been mainly an analogue of those in Bott [5]; while Bott has considered not only G^u/K and G^u/T, but also any compact kählerian homogeneous space $G^u/H = X^u$. Thus it will immediately call our attentions to the quotient manifold $Y_H = \Gamma\backslash G/H$ of $X_H = G/H$ by Γ, *in case* H *is contained in* K. Indeed, we may construct at once the commutative diagram analogous to that in § 3.

$$\begin{tikzcd}
Y_T \arrow[d, "\varpi_H"'] & X_T \arrow[l, "\phi_T"'] \arrow[r, "j_T"] \arrow[d, "\pi_H"'] & X_T{}^u \arrow[d, "\phi_H"'] \\
Y_H & X_H \arrow[l, "\phi_H"] \arrow[r, "j_H"'] & X_H{}^u
\end{tikzcd}$$

where the mapping j_H: $j_H(gH) = gU'$ $(X_H{}^u = G^c/U')$ is injective, since $U' \cap G = H$ is also valid. Further, Y_H being the total space of the holomorphic fibre bundle $Y_H(Y, K/H, K^c)$ with the kählerian homogeneous space K/H as the fibre, we infer from a theorem of Borel [9, p. 140] that Y_H is an algebraic manifold. Let $E'_\rho = G^c \times_{U'} \boldsymbol{C}^m$ be now the homogeneous vector bundle over $X_H{}^u$ associated to $\rho \in \mathrm{Hom}(H^c, GL(m, \boldsymbol{C})) \cong \mathrm{Hom}(U', GL(m, \boldsymbol{C}))$, then there will be defined the holomorphic vector bundle $\Gamma\backslash j_H{}^*E_\rho'$ over Y_H.

For such a bundle we can prove the vanishing theorems and the proportionality theorem in a quite similar way to the case $H=K$; though they would have at present no connections with automorphic forms in the usual sense.

Added in proof. The statement of the condition (b) for the discrete subgroup Γ in § 2 must be revised, because G acts on X not effectively in general: Denoting by Z the (finite) center of G (Z consists of all $g \in G$ which acts trivially on X), we have to replace Γ in (b) by $\Gamma/\Gamma \cap Z$.

College of General Education, University of Tokyo.

REFERENCES.

[1] A. Borel, "Les fonctions automorphes de plusieurs variables complexes," *Bulletin de la Société Mathématique de France*, vol. 80 (1952), pp. 167-182.

[2] ———, "Sur la cohomologie des espaces fibrés principaux et les espaces homogenes des groupes de Lie compacts," *Annals of Mathematics*, vol. 57 (1953), pp. 115-207.

[3] A. Borel and A. Weil, "Représentations linéaires et espaces homogènes kählериens des groupes de Lie compacts," *Seminaire Bourbaki*, May 1954. (Exposé by J-P. Serre.)

[4] A. Borel and F. Hirzebruch, "Characteristic classes and homogeneous spaces I," *American Journal of Mathematics*, vol. 80 (1958), pp. 459-538.

[5] R. Bott, "Homogeneous vector bundles," *Annals of Mathematics*, vol. 66 (1957), pp. 203-248.

[6] E. Calabi and E. Vesentini, "On compact, locally symmetric Kähler manifolds," *Annals of Mathematics*, vol. 71 (1960), pp. 472-507.

[7] H. Cartan, "Formes modulaires," *Seminaire H. Cartan*, 1957-1958, Exp. IV.

[8] Harish-Chandra, "Representations of semi-simple Lie groups on a Banach space VI," *American Journal of Mathematics*, vol. 78 (1956), pp. 564-628.

[9] F. Hirzebruch, *Neue topologische Methoden in der algebraischen Geometrie*, Berlin, Springer, 1956.

[10] ———, "Automorphe Formen und der Satz von Riemann-Roch," *Symposium International de Topologia Algebraica*, Mexico, 1958, pp. 129-144.

[11] ———, "Characteristic numbers of homogeneous domains," *Seminars on Analytic Functions*, vol. 2, Princeton, 1957, pp. 92-104.

[12] M. Ise, "Some properties of complex analytic vector bundles over compact homogeneous spaces," *Osaka Mathematical Journal*, vol. 12 (1960), pp. 217-252.

[13] N. Iwahori, *Theory of Lie Groups II* (in Japanese), Tokyô, Iwanami, 1958.

[14] K. Kodaira, "On a differential-geometric method in the theory of analytic stacks," *Proceedings of the National Academy of Sciences, U.S.A.*, vol. 39 (1953), pp. 1268-1273.

[15] ———, "On Kähler varieties of restricted type," *Annals of Mathematics*, vol. 60 (1954), pp. 28-48.

[16] Y. Matsushima and S. Murakami, "On vector bundle valued harmonic forms and automorphic forms on symmetric riemannian manifolds," to appear in *Annals of Mathematics*.

HARMONIC MAPPINGS OF RIEMANNIAN MANIFOLDS.*

By JAMES EELLS, JR. and J. H. SAMPSON.

Introduction. With any smooth mapping of one Riemannian manifold into another it is possible to associate a variety of invariantly defined functionals. Each such functional of course determines a class of extremal mappings, in the sense of the calculus of variations, and those extremals, in the very special cases thus far considered, play an important rôle in a number of familiar differential-geometric theories.

The present paper is devoted to a rather general study of a functional E of geometrical and physical interest, analogous to energy. Our central problem is that of deforming a given mapping into an extremal of E. Following an infinite-dimensional analogue of the Morse critical point theory, we construct gradient lines of E (in a suitable function space); and E is a decreasing function along those lines. With suitable metric and curvature assumptions on the target manifold (assumptions which cannot be entirely circumvented, in view of the examples of §§ 4E and 10D), we prove that the gradient lines do in fact lead to extremals (see Theorem 11A).

If $f: M \rightarrow M'$ is a smooth mapping of manifolds whose metrics are $g_{ij}dx^i dx^j$ resp. $g_{\alpha\beta}' dy^\alpha dy^\beta$, then the energy $E(f)$ is defined by the integral

$$E(f) = \tfrac{1}{2} \int_M g_{\alpha\beta}' \frac{\partial f^\alpha}{\partial x^i} \frac{\partial f^\beta}{\partial x^j} g^{ij} * 1,$$

where the f^α are local coordinates of the point $f(x)$, $*1$ being the volume element of M (assumed compact). Thus $E(f)$ can be considered as a generalization of the classical integral of Dirichlet. The Euler-Lagrange equations for E are a system of non-linear partial differential equations of elliptic type:

$$\Delta f^\alpha + \Gamma_{\beta\gamma}'^{\alpha} \frac{\partial f^\beta}{\partial x^i} \frac{\partial f^\gamma}{\partial x^j} \quad g^{ij} = 0;$$

Δ is the Laplace-Beltrami operator on M and the $\Gamma_{\alpha\gamma}'^{\alpha}$ are the Christoffel symbols on M'. Although this system is suggestive of the simple equation $\Delta u + \phi(u) \cdot \mathrm{grad}^2 u = 0$ for one unknown, there is in general very little connection between the two because of the phenomenon of curvature.

Received April 22, 1963.

* This work was supported in part by the National Science Foundation and the Office of Naval Research.

It has been necessary to go into the question of existence of solutions in rather great detail, owing to the want of general results for non-linear systems. Direct methods of the calculus of variations seem to lead to severe difficulties, and that is one reason why we have preferred to approach the problem through the gradient-line technique, which amounts to replacing the equations above by a system of parabolic equations whose relation to the elliptic system is analogous to that of Fourier's equation to Laplace's equation. This approach is of independent interest, in any case. Our methods are strongly potential-theoretic in nature. The local equations are first replaced by global equations of essentially the same form, embedding M' in a Euclidean space. A stability theorem is established showing that a solution of the resulting parabolic system does in fact produce a 1-parameter family of mappings of M into M'. Fundamental solutions of Laplace's equation and the heat equation on (compact) manifolds are used to establish *a priori* derivatives estimates and to construct solutions of the parabolic system, the latter being translated into a system of non-linear integro-differential equations of the Volterra type, following the method used by Milgram and Rosenbloom in a linear problem. Curvature enters in a manner not unlike that exploited by Bochner in [4].

Special cases of our extremal mappings go back to the very beginning of differential geometry. E.g. they include geodesics, harmonic functions, and minimal submanifolds. For minimal surfaces they were first studied locally by Bochner [2], in an explicitly Riemannian context. That work was carried to completion by Morrey [19]. In a report prepared by J. H. Sampson in February of 1954 at the Massachusetts Institute of Technology, the subject was taken up from a somewhat different point of view, and other geometrical applications were discussed. Since firm existence proofs were not then available, general publication of the results did not seem warranted. Shortly thereafter, J. Nash and, independently, F. B. Fuller [10] advanced the same definition as that on which this article is based, and Fuller described several examples. The problem has also been considered by E. Rauch.

The contents of the present paper are presented in the following order:

Chapter I. The concepts of energy and tension.
1. The energy integral
2. The tension field
3. Invariant formulation
4. Examples
5. The composition of maps

Chapter II. Deformations of maps.
6. Deformations by the heat equation
7. Global equations
8. Derivative bounds for the elliptic case
9. Bounds for the parabolic case
10. Successive approximations
11. Harmonic mappings

Added in proof: The theory of the energy functional (and its harmonic extremals) is the first-order case of a general theory of p-th order energy (and its polyharmonic extremals). See J. Eells, Jr. and J. H. Sampson, *L'énergie et les déformations en géometrie différentielle,* Colloque du CNRS (Proceding of a conference held in Grenoble in July, 1963) for a general formulation in fibre bundles.

Chapter I. The concepts of energy and tension.

1. The energy integral.

(A) Let M and M' denote complete Riemannian manifolds of dimensions n and m, respectively, and suppose further that M is closed (i.e., compact and without boundary) and oriented. In the interests of simplicity we assume that both manifolds and their Riemannian metrics are smooth (i.e., of class C^∞); however, it is not difficult to make minor modification to include differentiability class C^5. We will let $(x^1,\cdots,x^n)$ denote local coordinates on M in a neighborhood of a point P (said to be *centered at* P if $(0,\cdots,0)$ are its local coordinates), and $(y^1,\cdots,y^m)$ local coordinates on M'. Thus we can write the Reimannian metrics g and g' in these local coordinates as

$$ds^2 = g_{ij}dx^i dx^j, \qquad ds'^2 = g'_{\alpha\beta}dy^\alpha dy^\beta,$$

where we observe the summation convention; generally, when using local tensor calculus we follow the notations of Eisenhart [8]. We will denote covariant derivatives without the usual commas; e.g.,

$$f_i{}^\alpha = \partial f^\alpha/\partial x^i \qquad f_{ij}{}^\alpha = \partial^2 f^\alpha/\partial x^i \partial x^j - \Gamma_{ij}^k f_k{}^\alpha, \text{ etc.},$$

where Γ_{ij}^k denote the Christoffel symbols.

With any smooth map $f: M \rightarrow M'$ we assign a real number called its *energy,* as follows: For each point $P \in M$ we let $\langle\,,\rangle_P$ denote the inner product on the space of 2-covariant tensors of the tangent space $M(P)$ to M at P; thus if $(e^i)_{1\leq i\leq n}$ is a base for the cotangent space $M'(P)$ of $M(P)$ and $\alpha = \alpha_{ij}e^i \otimes e^j$ and α' are 2-covariant tensors of $M(P)$, then

$$\langle \alpha, \alpha' \rangle_P = \alpha_{ij}\alpha'_{pq}g^{ip}g^{jq},$$

where $g^{ik}g_{kj} = \delta_j^i$. Since the metric tensor g' of M' is 2-covariant, it induces through f a 2-covariant tensor field f^*g' on M, whence we can define the function $P \rightarrow \langle g(P), (f^*g')(P) \rangle_P$ on M. We will call $e(f)P = \frac{1}{2}\langle g(P), (f^*g')(P) \rangle_P$ the *energy density of f at P*. Its dual differential n-form $e(f) * 1$ can then be integrated over M, and with an eye toward the physical concept of kinetic energy $(mv^2/2)$ we define *the energy of the map f* by

$$E(f) = \int_M e(f) * 1. \tag{1}$$

In a local coordinate representation we have

$$E(f) = \tfrac{1}{2} \int_M g^{ij} f_i^\alpha f_j^\beta g'_{\alpha\beta} [\det(g_{ij})]^{\frac{1}{2}}\, dx^1 \wedge \cdots \wedge dx^n,$$

where $f_i^\alpha = \partial f^\alpha / dx^i$. Observe that if the local coordinates centered at P and $f(P)$ are both chosen to be locally Euclidean at their centers, then

$$e(f)P = \tfrac{1}{2} \sum_{i,\alpha} [f_i^\alpha(P)]^2,$$

so that $e(f)$ is *non-negative, and $E(f)$ vanishes when and only when f is a constant map.*

(B) Although the present work is devoted primarily to the functional E and its extremals, there will be the indications (e. g. the example § 4E) that we will want ultimately to consider other types of energy of maps. We give here a general method of constructing these.

Starting with the manifold M and any smooth symmetric 2-covariant tensor field α on M, we fix a point $P \in M$ and form the proper values of α relative to the metric tensor g of M; i. e., the n real roots of the equation $\det(g_{ij}(P)\lambda - \alpha_{ij}(P)) = 0$. Apart from their order, these proper values are intrinsically associated with α and P; thus we are led to forming their symmetric functions.

Definition. Let σ be any symmetric real function of n variables. For any symmetric 2-covariant tensor field α on M we let $\sigma(\alpha) : M \rightarrow R$ be the function such that $\sigma(\alpha)P =$ the σ-function of the proper values of $\alpha(P)$ relative to $g(P)$. The *σ-integral* of α is the number

$$I_\sigma(\alpha) = \int_M \sigma(\alpha) * 1.$$

In particular, let σ_p denote the p-th elementary symmetric function.

Then setting $\mu_0 = 1$, $\mu_p = \sigma_p / \binom{n}{p}$ we have Newton's inequalities $\mu_{p-1}\mu_{p+1} \leqq (\mu_p)^2$, with equality if and only if all proper values λ_i are equal; furthermore, if all $\lambda_i \geqq 0$, then

$$\mu_1 \geqq [\mu_2]^{1/2} \geqq [\mu_3]^{1/3} \geqq \cdots \geqq [\mu_n]^{1/n}, \text{ and} \tag{2}$$

$\mu_1/\mu_0 \geqq \mu_2/\mu_1 \geqq \cdots \geqq \mu_n/\mu_{n-1}$. These provide some sort of comparability of the various σ-integrals of α. We have

$$\sigma_1(\alpha)P = g^{ij}(P)\alpha_{ij}(P), \quad \sigma_n(\alpha)P = \det[\alpha_{ij}(P)]/\det[g_{ij}(P)],$$

and in general

$$\sigma_p(\alpha)P = (\det[g_{ij}])^{-1} \cdot \sum_{i_1 < \cdots < i_p} \begin{vmatrix} g_{11} \cdots g_{1n} \\ \alpha_{i_1 1} \cdots \alpha_{i_1 n} \\ \alpha_{i_p 1} \cdots \alpha_{i_p n} \\ g_{n1} \cdots g_{nn} \end{vmatrix}.$$

Remark. These integrals provide many variational problems of differential geometric interest. For example, if we take for α the Ricci tensor of M, then its proper values are the principal curvatures at a point, and $\sigma_1(\alpha)$ is the scalar curvature function. If we take for α a second fundamental form (for a given direction) of an immersion of M, then $\sigma_n(\alpha)$ is the Gauss-Kronecker curvature function (for that direction).

For the present purposes we take $\alpha = f^*g'$ for some map f. Then $(f^*g')_{ij} = f_i^\alpha f_j^\beta g'_{\alpha\beta}$, and $\sigma_1(f^*g') = \langle g, f^*g' \rangle$; we are led to the

Definition. For any map $f\colon M \rightarrow M'$ and any symmetric function σ of n variables, the *σ-energy of f* is the number

$$E_\sigma(f) = \int_M \sigma(f^*g') * 1.$$

We observe that *the energy of f is the $(\sigma_1/2)$-energy of f.*

Let $J_f(P)$ denote the Jacobian of f at P; i.e., the image under f of the unit n-vector of $M(P)$ defining the given orientation of M; thus

$$\sigma_n(f^*g')P = \langle J_f(P), J_f(P) \rangle / \det(g_{ij}(P)),$$

where the inner product is that of the space of n-vectors of $M'(f(P))$. The *volume of f* is the number

$$V(f) = \int_M [\sigma_n(f^*g')]^{\frac{1}{2}} * 1. \tag{3}$$

Note that $V(f) = 0$ if $n > m$.

Remark. If we consider $g + \lambda f^* g' = ds_\lambda^2$ as a perturbation of the metric of M for $\lambda \geqq 0$, then the corresponding volume element is

$$*1_\lambda = [\det(g_{ij} + \lambda f_i^\alpha f_j^\beta g'_{\alpha\beta})]^{\frac{1}{2}}\, dx^1 \wedge \cdots \wedge dx^n.$$

A simple calculation shows that

$$E(f) = \frac{d}{d\lambda} V_\lambda(M)_{\lambda=0},$$

where $V_\lambda(M) = \int_M *1_\lambda$.

(C) We give here an interpretation of $E(f)$ as a measurement of dispersion; the calculations of this paragraph will not be used in any essential way later.

Let $r'(P', Q')$ denote the geodesic distance between the points $P', Q' \in M'$, and suppose $P, Q \in M$ are points such that $f(P) = P'$, $f(Q) = Q'$. For a fixed P we consider the function $Q \to r'^2(P', Q')$; it is elementary that, for Q sufficiently near P, that function is smooth.

Let $\Delta f = g^{ij} f_{ij}$ denote the Laplace-Beltrami operator on the function f. We calculate the Laplacean of r'^2 *qua* function of Q. In coordinates centered at P and P' we have

$$\Delta_Q r'^2(P', Q') = g^{ij} \frac{\partial^2 r'^2}{\partial x^i \partial x^j} - g^{ij} \Gamma_{ij}^k \frac{\partial^2 r'^2}{\partial x^k}.$$

A simple calculation gives

$$\Delta_Q r'^2 = g^{ij} \frac{\partial^2 r'^2}{\partial y^\alpha \partial y^\beta} \frac{\partial y^\alpha}{\partial x^i} \frac{\partial y^\beta}{\partial x^j} + \frac{\partial r'^2}{\partial y^\alpha} \Delta y^\alpha.$$

Now

$$\frac{\partial r'^2}{\partial y^\alpha} = O(r'), \qquad \frac{\partial^2 r'^2}{\partial y^\alpha \partial y^\beta} = 2 g'_{\alpha\beta}(P') + O(r')$$

as $Q \to P$, whence

$$\Delta_Q r'^2(P', Q') = e(f)P + O(r').$$

For any smooth function u defined in a neighborhood of P let $\bar{u}_\epsilon(P)$ denote the average value ($\epsilon > 0$)

$$\bar{u}_\epsilon(P) = \int_{B_\epsilon} u*1/V(B_\epsilon),$$

where $B_\epsilon = \{Q \in M : f(P, Q) \leqq \epsilon\}$.

We require Maxwell's relation [16, p. 31. The formula is easily proved by expanding u in a Taylor's series at the center of a normal coordinate system.]:

$$\tilde{u}_\epsilon(P) = u(P) + \frac{\epsilon^2}{2(n+2)}\Delta u(P) + O(\epsilon^3).$$

Applying this to $u(Q) = r'^2(P', Q')$ and noting that $u(P) = 0$ and

$$\Delta u(P) = \lim_{Q\to P} \Delta_Q r'^2(P', Q') = e(f)P,$$

we obtain

$$\int_{B_\epsilon} r'^2(P', Q') {*1_Q}/V(B_\epsilon) = \frac{2\epsilon^2}{n+2} e(f)P + O(\epsilon^3).$$

If $\phi_\epsilon(P, Q)$ is a smooth function which is zero for $r(P, Q) \geqq \epsilon$ and equals $1/V(B_\epsilon)$ for $r(P, Q) \leqq \epsilon$, then we obtain from our last equation

$$\int_M \int_M r'^2(P', Q')\phi_\epsilon(P, Q) {*1} = \frac{2\epsilon^2}{(n+2)} E(f) + O(\epsilon^3).$$

Naturally certain uniformity conditions must be fulfilled by the error terms for the validity of this formula, but since we make no essential use of the result, we shall not insist on that point. *The left member of the preceding equation represents the mean square infinitesimal dispersion of the image points on M produced by f, and that quantity is estimated by the energy of f.*

2. The tension field.

(A) In this section we examine the extremals of E, interpreted as the zeros of the Euler-Lagrange operator associated with E. For this purpose we let $\pi\colon \mathcal{T}(M') \to M'$ denote the tangent vector bundle. We let $\mathcal{H}(M, M')$ denote the totality of smooth maps from M to M'; then for every $f \in \mathcal{H}(M, M')$ the set of smooth maps $u\colon M \to \mathcal{T}(M')$ such that $\pi \circ u = f$ forms a vector space $\mathcal{H}(f)$ with algebraic operations defined pointwise; such a u is called a *vector field along f*. We define an inner product in $\mathcal{H}(f)$ by

$$\langle u, v\rangle_f = \int_M \langle u(P), v(P)\rangle_{f(P)} {*1}.$$

For any vector field v along f the directional derivative of E in the direction v,

$$\nabla_v E(f) = \frac{d}{dt}[E(f_t)]_{t=0}, \text{ where } f_t(P) = \exp_{f(P)}(tv(P)),$$

is the endpoint of the geodesic segment in M' starting at $f(P)$ and determined in length and direction by the vector $tv(P) \in M'(f(P))$. We show in (B) below that *the Euler-Lagrange operator applied to a map f defines a vector field $\tau(f)$ along f which is the contravariant representative of the differential of E at f*; i. e.,

$$\nabla_v E(f) = -\langle \tau(f), v \rangle_f \text{ for all } v \in \mathfrak{H}(f).$$

Thus the maps f for which $\tau(f) = 0$ are (so to speak) the critical points of E.

(B) LEMMA. *Let $f_t \colon M \to M'$ be a smooth family of maps for t in some time interval $t_0 < t < t_1$. Then*

$$\frac{d}{dt} E(f_t) = -\int_M [\Delta f_t^\gamma + g^{ij} \Gamma'_{\alpha\beta}{}^\gamma f_{t,i}{}^\alpha f_{t,j}{}^\beta] g'_{\gamma\nu} \frac{\partial f_t{}^\nu}{\partial t} *1, \tag{4}$$

where Δ denotes the Laplace-Beltrami operator.

Proof. From the definition (1) we obtain for any $t \in (t_0, t_1)$

$$\frac{d}{dt} E(f_t) = \tfrac{1}{2} \int_M g^{ij} [2 f_{t,i}{}^\alpha \frac{\partial^2 f_t{}^\beta}{\partial x^j \partial t} g'_{\alpha\beta} + f_{t,i}{}^\alpha f_{t,j}{}^\beta \frac{\partial g'_{\alpha\beta}}{\partial y^\nu} \frac{\partial f_t{}^\nu}{\partial t}] *1.$$

The quantities

$$\xi^j = g^{ij} \frac{\partial f_t{}^\alpha}{\partial x^i} \frac{\partial f_t{}^\beta}{\partial t} g'_{\alpha\beta}$$

are the components of a contravariant vector field ξ on M, whence by Green's divergence theorem

$$\int_M \xi^j{}_j *1 = 0.$$

I. e.,

$$\int_M g^{ij} [f_{t,ij}{}^\alpha \frac{\partial f_t{}^\beta}{\partial t} g'_{\alpha\beta} + f_{t,i}{}^\alpha \frac{\partial^2 f_t{}^\beta}{\partial x^j \partial t} g'_{\alpha\beta} + f_{t,i}{}^\alpha \frac{\partial f_t{}^\beta}{\partial t} \frac{\partial g'_{\alpha\beta}}{\partial y^\nu} \frac{\partial f_t{}^\nu}{\partial x^j}] *1 = 0.$$

A routine calculation now gives (4).

For any smooth map $f \in \mathfrak{H}(M, M')$ we set

$$\tau(f)^\gamma(P) = \Delta f^\gamma(P) + g^{ij}(P) \Gamma'_{\alpha\beta}{}^\gamma(f(P)) f_i{}^\alpha(P) f_j{}^\beta(P). \tag{5}$$

It is clear that $\tau(f)^\gamma$ is unaffected by any transformation of the local coordinates on M near P, and for any such transformation of coordinates near $f(P)$ we see that $\tau(f)^\gamma(P)$ transforms as a contravariant vector in $M'(f(P))$; see § 3 for an invariant formulation. Thus with every $f \in \mathfrak{H}(M, M')$ we have a variation $\tau(f) \in \mathfrak{H}(f)$.

Definition. We call $\tau(f)$ *the tension field of the map* f, and we say that *f is a harmonic map if $\tau(f) = 0$.* Thus (4) becomes

$$\frac{d}{dt} E(f_t) = -\int_M \langle \tau(f), \frac{df_t}{dt} \rangle *1.$$

Suppose now that f is an extremal (with respect to small deformations) of E. If we apply (4) to suitably chosen local deformations (e.g., confined to small geodesic balls on M), we obtain from a standard argument that *the Euler-Lagrange equation for the energy functional E is $\tau(f)=0$. In local coordinates that equation is elliptic.*

Remark. It is a simple matter to modify these constructions to include the case that M has a boundary; we should require in Lemma 2B that f_t is a constant function of t on the boundary.

PROPOSITION. *Every map $f\colon M\rightarrow M'$ of class C^2 which satisfies $\tau(f)=0$ is smooth. If M and M' are both analytic Riemannian manifolds, then every such map is analytic.*

The proof of the first assertion is easily obtained by inductive application of [14, Theorem 3, p. 210]; analyticity follows directly from [22, p. 4].

(C) Our next step is to interpret the tension field $\tau(f)$ in terms of special coordinate systems on M and M'.

For any $P\in M$ we let $\exp_P\colon M(P)\rightarrow M$ be that map which assigns to each vector $u\in M(P)$ the end point of the geodesic segment emanating from P and determined in length and direction by u; it is elementary that $\exp_P$ is a smooth surjective map, carrying a neighborhood of $0\in M(P)$ diffeomorphically onto a neighborhood V of P in M. If we now choose an orthonormal frame in $M(P)$, then we can use the inverse of $\exp_P$ to refer each point $Q\in V$ to the components relative to that frame. These are called *normal coordinates in V centered at P,* and they form a coordinate system admissible for the differentiable structure of M.

In terms of P-centered normal coordinates we have

1) the metric tensor $g_{ij}(P)=\delta_{ij}$, since the frame is orthonormal;

2) the Christoffel symbols $\Gamma_{ij}^k(P)=0$ for all i, j, k; that is because the equations of the geodesics through P are linear, whence $\Gamma_{ij}^k(P)u^iu^j=0$ for all vectors $u\in M(P)$; see Eisenhart [8, p. 54]. In particular, $f_{ij}^\gamma(P)=\partial^2f^\gamma(P)/\partial x^i\partial x^j$, from which we obtain the

LEMMA. *For any $f\in\mathscr{H}(M,M')$ and point $P\in M$ let us fix normal coordinate systems V and V' centered at P and $P'=f(P)$, respectively; in terms of these we have*

$$\tau(f)^\gamma(P)=\sum_{i=1}^n\frac{\partial^2f^\gamma(P)}{\partial(x^i)^2}. \qquad (1\leqq\gamma\leqq m)$$

Thus f is harmonic if and only if at each pair of points P, $f(P)$ there are such coordinates in which f satisfies the Euclidean-Laplace equation at P.

Example. Suppose M' is flat; i.e., its Riemannian curvature is zero at every point. It is well known that then M' admits a smooth coordinate covering such that in each coordinate chart we have $\Gamma'_{\alpha\beta}{}^{\gamma} \equiv 0$. Then $\tau(f)^{\gamma} = \Delta f^{\gamma}$, and in such local coordinates *the equation* $\tau(f) = 0$ *is linear.* In particular, if M' is the Euclidean space $\boldsymbol{R}^m$, then a map $f\colon M \to \boldsymbol{R}^m$ is harmonic if and only if it is constant, by the maximum principle.

Example. If M and M' are Lie groups with bi-invariant Riemann metrics, and if $f\colon M \to M'$ is a homomorphism, then f is harmonic. It suffices to verify this at the identity of M; taking canonical coordinates (which are normal; see Chevalley [5, p. 118]) at the neutral elements of M and M', we see that the representation of f is linear, whence $\tau(f) = 0$ by Lemma 2C.

Example. If M and M' are Kähler manifolds and $f\colon M \to M'$ is a holomorphic map, then f is harmonic relative to the associated real Riemann structures on M and M'. Namely, we take local holomorphic coordinates $z^j = x^j + \sqrt{-1}\, x^{h+j}$ and $w^{\alpha} = y^{\alpha} + \sqrt{-1}\, y^{k+\alpha}$ (where $\dim M = 2h$, $\dim M' = 2k$) centered at some points $P \in M$, $f(P) \in M'$. Because the w^{α} are holomorphic functions of the z^j and because M is Kähler, they satisfy Laplace's equation $\Delta w^{\alpha} = 0$. Now Δ is a real operator, which we have $\Delta y^{\alpha} = 0 = \Delta y^{k+\alpha}$ for $1 \leqq \alpha \leqq k$. Since M' is Kähler, the w^{α} can be chosen to be normal coordinates, whence all $\Gamma'_{\alpha\beta}{}^{\gamma} = 0$ at $f(P)$. Then $\tau(f)P = 0$.

(D) Suppose *$f\colon M \to M'$ is a Riemannian immersion.* Thus for each $P \in M$ the differential $f_*(P)$ of f at P maps $M(P)$ isometrically into $M'(f(P))$; i.e., $g = f^*g'$. If $\xi_{\sigma|}$ $(1 \leqq \sigma \leqq m - n)$ is an orthonormal frame in (some coordinate system) on M' orthogonal to M, then

$$b_{\sigma|ij} = (f_{ij}{}^{\gamma} + \Gamma'_{\alpha\beta}{}^{\gamma} f_i{}^{\alpha} f_j{}^{\beta})\, g'_{\gamma\nu}\, \xi_{\sigma}{}^{\nu}{}_{|} = b_{\sigma|ji} \tag{6}$$

are the components of the second fundamental form $\beta(f)$ on M in M' relative to that frame (Eisenhart [8, § 50]). The vector field ξ along the map f defined by

$$\xi(P) = \sum_{i=1}^{m-n} g^{ij}(P)\, b_{\sigma|ij}(P)\, \xi_{\sigma|}(P) \tag{7}$$

for all $P \in M$ is independent of choice of frame. It is traditionally called *the mean normal curvature field of the immersion.* The following formula is a well known interpretation:

$$\xi(P) = -\sum_{\sigma=1}^{m-n} \operatorname{div}(\xi_{\sigma|}(P))\, \xi_{\sigma|}(P).$$

Note that the coefficient of each $\xi_{\sigma|}$ is the first elementary symmetric function of the proper values of $b_{\sigma|}$ relative to g; that indicates how the following results (and those of § 4C) will generalize when we replace energy by σ-energy.

A Riemannian immersion is said to be a *minimal immersion* if $\xi = 0$ on M. The minimal immersions are precisely those which are the extremals of the volume functional V (Eisenhart [8, p. 176]). Usually the notion of minimal immersion is taken in a somewhat broader context, as a smooth map which is an extremal of V. There are examples (even for $M' = \mathbf{R}^m$ and $n = 2$) of analytic extremals of V which are not strictly immersions, for their Jacobians vanish at isolated points.

From (7) we conclude that

$$\begin{aligned}\xi &= \sum_{\sigma=1}^{m-n} g^{ij}(f_{ij}{}^{\gamma} + \Gamma'_{\alpha\beta}{}^{\gamma} f_i{}^{\alpha} f_j{}^{\beta}) g'_{\gamma\nu} \xi_{\sigma|}{}^{\nu} \xi_{\sigma|} \\ &- \sum_{\sigma=1}^{m-n} \langle \tau(f), \xi_{\sigma|} \rangle \xi_{\sigma|}.\end{aligned}$$

But $(f_{ij}{}^{\gamma} + \Gamma'_{\alpha\beta}{}^{\gamma} f_i{}^{\alpha} f_j{}^{\beta}) g'_{\gamma\nu} f_k{}^{\nu} = 0$ $(1 \leqq k \leqq n)$ in general (Eisenhart [8, p. 160]); i.e., $\tau(f)(P)$ *is perpendicular to* $M(P)$.

PROPOSITION. *Let $f\colon M \to M'$ be a Riemannian immersion. Then $e(f) = n/2$. The tension field $\tau(f)$ coincides with the mean normal curvature field on f. In particular, f is harmonic if and only if it is minimal.*

Note that any isometry $f\colon M \to M$ is harmonic, and that any covering map is harmonic.

Remark. For any $u \in M'(f(P))$ perpendicular to $M(P)$ we have $\langle \tau(f), u \rangle = \operatorname{trace}(\beta(f)u)$.

(E) For any map $f \in \mathfrak{H}(M, M')$ we define $F\colon M \to M \times M'$ by $F(P) = (P, f(P))$; then F is a smooth imbedding, but not isometric in the product metric on $M \times M'$. The following result is immediate:

PROPOSITION. *For any $f \in \mathfrak{H}(M, M')$ we have $e(F) = n/2 + e(f)$. Furthermore, f is harmonic if and only if F is harmonic.*

3. Invariant formulation.

(A) In this section we express our problem in terms of differential forms with values in a vector bundle. It turns out that in some sense the harmonic theory of such forms gives us an infinitesimal solution to our problem. We begin by summarizing the general theory; for details see Spencer [27] or Bochner [4].

1. Let $W \rightarrow M$ be a smooth vector bundle over M with fibre dimension m. If $\mathcal{T}^{[p]}(M) \rightarrow M$ denotes the bundle of p-covectors of M and if $W \otimes \mathcal{T}^{[p]}(M) \rightarrow M$ is the tensor product bundle, then the smooth sections of that latter bundle are called the *smooth p-forms on M with values in W*; their totality forms a vector space, which we denote by $A^p(M, W)$.

2. Suppose that W has a Riemannian structure; i. e., a given reduction of its structural group to the orthogonal group in dimension m. If $W \rightarrow W^*$. denotes the dual bundle of W, then we have a bundle isomorphism $\psi\colon W \rightarrow W^*$. Taken together with the canonical isomorphism $*\colon \mathcal{T}^{[p]}(M) \rightarrow \mathcal{T}^{[n-p]}(M)$ induced from the given Riemannian structure on M, we have the natural map $\psi \otimes *\colon A^p(M, W) \rightarrow A^{n-p}(M, W^*)$. In particular, the evaluation of W^* on W induces a bilinear pairing

$$\#\colon A^p(M, W) \times A^p(M, W) \rightarrow A^n(M),$$

the vector space of real valued n-forms on M. Thus $A^p(M, W)$ has the *inner product*

$$\langle \phi, \psi \rangle = \int_M \phi \# \psi. \tag{8}$$

3. Assume next that W has a connection which is compatible with its Riemannian structure; i. e., the covariant differential of the tensor field defining the orthogonal reduction of W is zero. We say that *W is a Riemannian-connected bundle*. Then we have a *linear map* $\bar{d}\colon A^p(M, W) \rightarrow A^{p+1}(M, W)$, which can be described as follows: If $\mathcal{U}$ is a locally finite covering of M by coordinate systems, then for each $U \in \mathcal{U}$ the connection in W can be given by a certain $m \times m$ matrix $\theta^U = (\theta_\alpha{}^{U\beta})$ of 1-forms in U. Similarly any form $\phi \in A^p(M, W)$ defines an m-tuple $\phi^U = (\phi_\beta{}^U)$ of p-forms in U. If d denotes the exterior differential operator in U, then we define $\bar{d}\phi$ by giving its representation in each U by the formula (In this section (3A) we violate our convention of lettering subscripts denote covariant differentiation. We adopt here a special symbol for that concept.)

$$\begin{aligned}(\bar{d}\phi)^U &= d\phi^U + \theta^U \wedge \phi^U \\ &= (d\phi_1{}^U + \theta_1{}^{U\beta} \wedge \phi_\beta{}^U, \cdots, d\phi_m{}^U + \theta_m{}^{U\beta} \wedge \phi_\beta{}^U).\end{aligned}$$

4. We can now develop the theory of harmonic forms on M with values in a Riemannian-connected bundle. We have the adjoint map $\bar{\delta}\colon A^p(M, W) \rightarrow A^{p-1}(M, W)$, characterized by the formula

$$\langle \bar{d}\phi, \psi \rangle = \langle \phi, \bar{\delta}\psi \rangle.$$

In particular, for any $\phi \in A^1(M, W)$ and any coordinate chart U we have 1-forms $\phi_\alpha{}^U$ $(1 \leqq \alpha \leqq m)$, which can be expressed

$$\phi_\alpha{}^U = \phi_{\alpha j}{}^U \, dx^j \text{ in the coordinates } (x^1, \cdots, x^n) \text{ in } U.$$

If we write similarly

$$\theta_\alpha{}^{U\beta} = T_{\alpha i}{}^\beta \, dx^i,$$

then

$$(9) \qquad \begin{aligned} (\bar{d}\phi)_\alpha{}^U &= \Big(\frac{\partial \phi_{\alpha j}{}^U}{\partial x^i} + T_{\alpha i}{}^\beta \phi_{\beta j}{}^U\Big) dx^i \wedge dx^j, \\ (\delta\phi)_\alpha{}^U &= -g^{ij}(\nabla_i \phi_{\alpha j}{}^U + T_{\alpha i}{}^\beta \phi_{\beta j}{}^U), \end{aligned}$$

letting ∇_i denote covariant differentiation on M relative x^i. The *Dirichlet integral of* $\phi \in A^p(M, W)$ is $2\bar{D}(\phi) = \langle \bar{d}\phi, \bar{d}\phi \rangle + \langle \bar{\delta}\phi, \bar{\delta}\phi \rangle$, and its associated *Laplace operator is* $\bar{\Delta} = -(\bar{d}\bar{\delta} + \bar{\delta}\bar{d})$. *A form* $\phi \in A^p(M, W)$ *is harmonic if* $\bar{\Delta}\phi = 0$; that condition is equivalent to the pair $\bar{d}\phi = 0 = \bar{\delta}\phi$.

It is known that a Green's form exists for the operator $\bar{\Delta}$, from which it can be proved that there is a decomposition

$$(10) \qquad A^p(M, W) = H^p(M, W) \oplus [\bar{d}A^{p-1}(M, W) + \bar{\delta}A^{p+1}(M, W)],$$

where $H^p(M, W)$, the space of *harmonic p-forms with values in* W, is orthogonal to the other two summands.

5. The curvature of the connection can be expressed in U by an $m \times m$ matrix of 2-forms $\Theta^U = (\Theta_{\alpha\beta}{}^U)$ where $\Theta^U = d\theta^U + \theta^U \wedge \theta^U$. If ∇ denotes the covariant differential of M, then we define *the covariant differential* $\bar{\nabla}$ on $\phi \in A^p(M, W)$ by

$$\bar{\nabla}^U \phi^U = \nabla \phi^U + \theta^U \wedge \phi^U.$$

Letting $R_i{}^h$ denote $R_i{}^h I$, the diagonal $m \times m$ matrix each of whose diagonal terms is the Ricci tensor field $R_i{}^h$ (we follow the sign convensions in Eisenhart in defining $R_{ij} = R^h{}_{ijh}$), we have for any $\phi \in A^1(M, W)$ the following expression for the components of the Laplacean of ϕ in U:

$$(-\bar{\Delta}\phi)_i{}^U = -(\bar{\nabla}^U)^k(\bar{\nabla}^U)_k \phi_i{}^U + {}^t[(\Theta^U)_i{}^h - R_i{}^h]\phi_h{}^U,$$

where t denotes the transposition of matrices.

If $\phi, \psi \in A^1(M, W)$, define the function $\phi \cdot \psi = *(\phi \# \psi)$ on M; then in U we have (usual Laplacean Δ of functions)

$$\begin{aligned} \Delta(\tfrac{1}{2}\phi \cdot \phi) = \phi^U \cdot (-\bar{\Delta}\phi^U) &- (\bar{\nabla}^U)_i(\phi^U)_k a^U (\bar{\nabla}^U)^i(\phi^U)^k \\ &- {}^t[(\Theta^U)_i{}^h - R_i{}^h](\phi^U)_h a^U (\phi^U)^i, \end{aligned}$$

where a^U denotes the Riemannian structure of W in U. For each $\phi \in A^1(M, W)$ define the function

$$Q(\phi) = {}^t[(\Theta^U)_i{}^h - R_i{}^h](\phi^U)_h a^U (\phi^U)^i. \tag{11}$$

The matrix (of functions) of Q is

$$Q_{\alpha\beta}{}^{hi} = g^{hk} g^{ij} [(a_{\alpha\gamma}{}^U \Theta_{\gamma\beta}{}^U)_{kj} - a_{\alpha\beta}{}^U R_{kj}].$$

We consider this as an $nm \times nm$ matrix in the subscripts (αh), (βi); as such, it is symmetric: $Q_{\alpha\beta}{}^{hi} = Q_{\beta\alpha}{}^{ih}$.

The integral over M of $\Delta(\phi \cdot \phi/2)$ is always zero, by Green's theorem. Thus *if* ϕ *is harmonic we have*

$$\int_M Q(\phi) * 1 = -\int_M (\bar{\nabla}^U)_i (\phi^U)_k a^U (\bar{\nabla}^U)^i (\phi^U)^k * 1 \leqq 0. \tag{12}$$

(B) Given any $f \in \mathcal{H}(M, M')$, let $f^{-1}\mathcal{T}(M') \to M$ be the induced vector bundle; it is clearly Riemannian-connected. Let us interpret the preceding development for that bundle.

First of all, the elements of $A^0(M, f^{-1}\mathcal{T}(M'))$ are canonically identified with the vector fields along f (i.e., with the elements of the space $\mathcal{H}(f)$, in the notation of § 2A). Secondly, for any $P \in M$ the differential $f_*(P)$: $M(P) \to M'(f(P))$ is a linear map, to be considered as an element of $M'(f(P)) \otimes M^{[1]}(P)$; otherwise said, the assignment $P \to f_*(P)$ determines a specific 1-form $f_* \in A^1(M, f^{-1}\mathcal{T}(M'))$. Thirdly, we have

$$E(f) = \tfrac{1}{2} \int_M f_* \# f_*.$$

LEMMA. *For any* $f \in \mathcal{H}(M, M')$ *we have* $\bar{d} f_* = 0$; *i.e.,* f_* *is orthogonal to* $\bar{\delta} A^2(M, f^{-1}\mathcal{T}(M'))$. *Thus* $\bar{\Delta} f_* = -\bar{d}\bar{\delta} f_*$.

Proof. Take a coordinate chart U on M, and write

$$\theta_\gamma{}^{U\beta} = \Gamma'{}_{\alpha\beta}{}^\gamma f_i{}^\alpha \, dx^i. \tag{13}$$

Then $(f_*{}^U)^\beta = f_j{}^\beta \, dx^j$, whence $(1 \leqq \alpha \leqq m)$

$$(\bar{d} f_*{}^U)_\gamma = \frac{\partial^2 f^\gamma}{\partial x^i \partial x^j} \, dx^i \wedge dx^j + \Gamma'{}_{\alpha\beta}{}^\gamma f_i{}^\alpha f_j{}^\beta \, dx^i \wedge dx^j,$$

which is zero, because both coefficients are symmetric in i, j.

Similarly, the variation $\bar{\delta} f_* \in A^0(M, f^{-1}\mathcal{T}(M'))$ has coordinate representation

$$(\bar{\delta} f_*)_\gamma{}^U = -g^{ij}\{\nabla_i \nabla_j f^\gamma + \Gamma'{}_{\alpha\beta}{}^\gamma f_i{}^\alpha f_j{}^\beta\} = -\tau(f)^\gamma.$$

PROPOSITION. *For any* $f \in \mathcal{H}(M, M')$ *its differential* f_* *is a closed* 1-*form. Its tension field* $\tau(f) = -\bar{\delta} f_*$, *the divergence of its differential. The map* f *is harmonic if and only if its differential is a harmonic* 1-*form.*

Definition. For any map $f \in \mathcal{H}(M, M')$ its *fundamental form* $\beta(f)$ is the covariant differential $\tilde{\nabla} f_*$ of its differential. Thus $\beta(f)$ is the $f^{-1}\mathcal{T}(M')$-valued 2-covariant tensor field on M whose coordinate representation is

$$f_{;ij}{}^{\gamma} = f_{ij}{}^{\gamma} + \Gamma'_{\alpha\beta}{}^{\gamma} f_i{}^{\alpha} f_j{}^{\beta} = f_{;ji}{}^{\gamma}.$$

The tension field $\tau(f)$ is just the trace of $\beta(f)$; i.e., $\tau^{\gamma}(f) = g^{ij} f_{;ij}{}^{\gamma}$ $(1 \leqq \gamma \leqq m)$. It follows from § 2D that if f is a Riemannian immersion, then $\beta(f)$ is the second fundamental form of M in M'. Analogously, let us say that a map $f \in \mathcal{H}(M, M')$ *is totally geodesic* if $\beta(f) = 0$ on M; we will see as a consequence of Corollary 5A below that *totally geodesic maps map geodesics into geodesics.*

(C) Let us consider the function $Q(f_*)$; from the expression $(\Theta_{\alpha\beta}{}^{U})_i{}^{h} = g^{hk} R'_{\alpha\beta\gamma\delta} f_k{}^{\gamma} f_i{}^{\delta}$ we compute (11), taking into account the skew symmetry $\Theta_{\alpha\beta}{}^{U} = -\Theta_{\beta\alpha}{}^{U}$ to obtain the

LEMMA. *For any smooth map* $f\colon M \to M'$ *we have*

$$Q(f_*) = -R'_{\alpha\beta\gamma\delta} f_i{}^{\alpha} f_j{}^{\beta} f_k{}^{\gamma} f_l{}^{\delta} g^{ik} g^{jl} - R^{ij} f_i{}^{\alpha} f_j{}^{\beta} g'_{\alpha\beta}. \tag{14}$$

Its matrix (for arbitrary forms $\phi \in A^1(M, f^{-1}\mathcal{T}(M'))$*) is*

$$Q_{\alpha\beta}{}^{ij} = -R'_{\alpha\beta\gamma\delta} f_k{}^{\gamma} f_l{}^{\delta} g^{ik} g^{jl} - R^{ij} g'_{\alpha\beta}. \tag{15}$$

If f *is harmonic, then*

$$\Delta e(f) = |\beta(f)|^2 + Q(f_*). \tag{16}$$

We will refer to the matrix (15) as *the Ricci transformation on the tensor product bundle* $f^{-1}\mathcal{T}(M') \otimes \mathcal{T}^{[1]}(M)$. Observe that if f is a real-valued function on M, then that Ricci transformation is just that given by the Ricci tensor of M.

Remark. The above computations can of course be made without passing through the medium of vector-bundle-valued differential forms. One starts by applying the Ricci identities (Eisenhart [8, p. 30]) to the direct evaluation of $\Delta e(f)$, and then reads off the appropriate terms.

The next result follows the well known pattern of Bochner; in [2] Bochner has also applied the method in a special case for maps.

THEOREM. *If* $f\colon M \to M'$ *is a harmonic map, then*

$$\int_M Q(f_*)*1 \leqq 0,$$

and equality holds when and only when f is totally geodesic. Furthermore, if $Q(f_*) \geqq 0$ on M, then f is totally geodesic and has constant energy density $e(f)$.

Proof. This follows at once from Stokes' Theorem

$$\int_M \Delta e(f)*1 = 0,$$

applied to (16); for if $\Delta e(f) \geqq 0$, then $\Delta e(f) = 0$ everywhere, whence $e(f)$ is a constant function.

Following the conventions of Eisenhart [8], we say that the Ricci curvature of M is non-negative if at every point $P \in M$ the matrix $(-R_{ij}(P))$ is positive semi-definite.

COROLLARY. *Suppose that the Ricci curvature of M is non-negative and that the Riemannian curvature of M' is non-positive. Then a map $f\colon M \to M'$ is harmonic if and only if it is totally geodesic. Furthermore,*

1) *if there is at least one point of M at which its Ricci curvature is positive, then every harmonic map $f\colon M \to M'$ is constant;*

2) *if the Riemannian curvature of M' is everywhere negative, then every harmonic map $f\colon M \to M'$ is either constant or maps M onto a closed geodesic of M'.*

Proof. The theorem shows that

$$R'_{\alpha\beta\gamma\delta} f_i{}^{\alpha} f_j{}^{\beta} f_k{}^{\gamma} f_l{}^{\delta} g^{ik} g^{jl} + R^{ij} f_i{}^{\alpha} f_j{}^{\beta} g'_{\alpha\beta} \equiv 0.$$

If hypothesis 1) is satisfied at $P \in M$, then $f_*(P) = 0$, whence the constant $e(f) = 0$; i.e., f is a constant map. If hypothesis 2) is satisfied at $P' = f(P) \in M'$ and we take normal coordinates centered at P, then the $f_*(P)$-image of the tangent space $M(P)$ has dimension $\leqq 1$. If it has dimension 0 at any $f(P)$, then again $e(f) = 0$; otherwise, the image $f_*(P)M(P)$ has constant dimension 1. Because f is totally geodesic, the conclusion follows.

Example. If $f\colon M \to M'$ is a harmonic immersion, then $e(f) = n/2$, and $|\beta(f)|^2 + Q(f_*) = 0$. This relation also follows from Gauss's equations (Eisenhart [8, p. 162])

$$R_{ijkl} - R'_{\alpha\beta\gamma\delta} f_i{}^{\alpha} f_j{}^{\beta} f_k{}^{\gamma} f_l{}^{\delta} = \sum_{\sigma=1}^{m-n} (b_{\sigma|ik} b_{\sigma|jl} - b_{\sigma|il} b_{\sigma|jk})$$

by multiplying by $g^{ik}g^{jl}$ and summing. We obtain

$$R^{ij}g_{ij} + R'_{\alpha\beta\gamma\delta}f_i^{\alpha}f_j^{\beta}f_k^{\gamma}f_l^{\delta}g^{ik}g^{jl} = |\beta(f)|^2 - \sum_{\sigma=1}^{m-n}\mu_\sigma^2,$$

where $\mu_\sigma = g^{ij}b_{\sigma|ij}$ is the σ-th component of the mean normal curvature of the immersion (see (7)); each $\mu_\sigma = 0$ if f is harmonic. For instance, *if the Riemannian curvature of M' is non-positive and if the scalar curvature $R = R^{ij}g_{ij}$ of M is negative at some point, then there is no harmonic immersion of M in M'.*

4. Examples.

(A) *The case* $\dim M = 1$. Let us take for M the unit circle S^1, coordinatized by the central angle θ. For any $f \in (S^1, M')$ we have

$$e(f) = \tfrac{1}{2}\frac{df^\alpha}{dt}\frac{df^\beta}{dt}g'_{\alpha\beta}, \text{ whence } E(f) = \tfrac{1}{2}\int_M \left|\frac{df}{dt}\right|^2 *1.$$

The tension field is

$$\tau(f)^\gamma(t) = \frac{d^2f^\gamma}{dt^2} + \Gamma'^{\gamma}_{\alpha\beta}\frac{df^\alpha}{dt}\frac{df^\beta}{dt},$$

which (when the parameter of f is proportional to arc length) is often called *the curvature* (or *acceleration*) *of f*. We have $\langle\tau(f), \frac{df}{dt}\rangle = 0$, and $\tau(f)$ is proportional to the geodesic curvature vector field along f. Then *f is harmonic if and only if f defines a closed geodesic on M'*. It is well known that *if M' is compact, then in every homotopy class of maps $S^1 \to M'$ there is a harmonic map* (and furthermore, one which minimizes the *length* in that homotopy class.) On the other hand, without further restrictions on g' there are complete Riemannian manifolds M' and non-trivial homotopy classes of maps $S^1 \to M'$ having no harmonic representatives; see § 10 below.

(B) *The case* $\dim M = 2$. We established here certain relations showing the close connection of our problem with the Plateau problem, in its potential theoretic formulation (Morrey [19] and Bochner [2]); incidentally, we see that in our energy theory the cases $\dim M \leq 2$ are favored.

Recall that a map $h\colon M \to M'$ *is conformal* if there is a smooth function $\theta\colon M \to R$ such that

$$h^*g' = \exp(2\theta)g.$$

Thus the differential h_* preserves orthogonality and dilatates uniformly. Clearly such a map is a smooth immersion, and has energy density $e(h) = n\exp(2\theta)/2$.

PROPOSITION. *If* $\dim M = n = 2$ *and* $h\colon M \to M$ *is a conformal diffeomorphism, then for all* $f \in \mathcal{H}(M,M')$ *we have* $E(f\circ h) = E(f)$. *Moreover,* h *is harmonic.*

Proof. First of all, $(f\circ h)_i^\alpha = f_p^\alpha h_i^p$, whence $2e(f\circ h) = g^{ij}h_i^p h_j^q f_p^\alpha f_q^\beta g'_{\alpha\beta}$. The conformaltity condition for h implies $g^{ij}h_i^p h_j^q = \exp(2\theta)g^{pq}$, and substituting gives

$$2e(f\circ h) = \exp(2\theta)g^{pq}f_p^\alpha f_q^\beta g'_{\alpha\beta} = 2\exp(2\theta)e(f).$$

Secondly, we have

$$\begin{aligned} h^*(*1) &= [\det(g_{pq}h_i^p h_j^q)/\det(g_{ij})]^{\frac{1}{2}} *1 \\ &= \exp(n\theta)*1, \end{aligned}$$

so that if $n = 2$ we have $h^*(e(f)*1) = e(f\circ h)*1$.

Finally, in suitable local coordinates on M we have

$$\Gamma_{ij}^k = \Gamma'^k_{ij} - (\delta_j^k\theta_i + \delta_i^k\theta_j - g^{kp}g_{ij}\theta_p),$$

and direct computation shows that

$$\tau(h)^i = (2-n)g^{ij}\theta_j \quad . \qquad\qquad (1 \leqq i \leqq n).$$

PROPOSITION. *If* $\dim M = 2$, *then for any* $f \in \mathcal{H}(M,M')$ *we have* $V(f) \leqq E(f)$. *Equality holds when and only when* f *is conformal.*

Proof. The first statement follows immediately from the inequalities (2). Suppose f is conformal; then for $n = 2$ we obtain

$$V(f) = \int_M \exp(2\theta)\,[g_{11}g_{22} - g_{12}{}^2]^{\frac{1}{2}}\,dx^1dx^2 = \int_M e(f)*1 = E(f).$$

Conversely, if $V(f) = E(f)$ we conclude that at every point of M

$$2[\det(f^*g')_{ij}]^{\frac{1}{2}} = g^{ij}(f^*g')_{ij}[\det(g_{ij})]^{\frac{1}{2}}.$$

In isothermal coordinates on M we have

$$[(f^*g')_{11} - (f^*g')_{22}]^2 = -\{(f^*g')_{12}\}^2,$$

whence $(f^*g')_{11} - (f^*g')_{22} = 0 = (f^*g')_{12}$. Defining $\theta\colon M \to R$ by

$$\exp(2\theta) = [\det(f^*g')_{ij}/\det(g_{ij})]^{\frac{1}{2}},$$

we obtain $\exp(2\theta)g_{ij} = (f^*g')_{ij}$.

COROLLARY. *If a map* $f_0\colon M \to M'$ *minimizes* V *and is conformal, then* f_0 *minimizes* E.

For any $f \in \mathfrak{H}(M,M')$ we have

$$E(f_0) = V(f_0) \leqq V(f) \leqq E(f).$$

Remark. If M is a Riemann surface, then its complex structure defines a conformal equivalence class of Riemann structures. The energy of any map $f\colon M \to M'$ therefore depends only on the complex structure of M.

(C) *Harmonic fibre maps.* For any map $f \in \mathfrak{H}(M,M')$ and $P \in M$ we have the vector space $M_V(P) = \{u \in M(P) : f_*(P)u = 0\}$; the vectors in $M_V(P)$ are called *vertical.* Let $M_H(P)$ denote its orthogonal complement in $M(P)$. Suppose that for all $P \in M$ the differential $f_*(P)$ maps $M_H(P)$ isometrically onto $M'(f(P))$. Then f is a locally trivial surjective fibre map (see Hermann [13]); in particular, f determines an almost product structure on M (i.e., the structural group O_n of $\mathfrak{F}(M)$ admits a reduction to the product group $O_m \times O_{n-m}$). We will call a map $f\colon M \to M'$ *a Riemannian fibration.*

Remark. There are smooth fibrations $f\colon M \to M'$ having no Lie structural group; e.g., there are non-trivial compact smooth fibrations over the 3-sphere S^3 (which cannot have a Lie structural group G, since $\pi_2(G) = 0$). It is a consequence of a theorem of Hermann [13] that the manifolds M, M' admit no Riemann structures compatible with f as above, for which the fibres are totally geodesic.

LEMMA. *If $f\colon M \to M'$ is a Riemannian fibration, then for any $P \in M$ there are coordinate charts U and U' centered at P and $f(P)$ respectively, in terms of which $f_i^\alpha(P) = \delta_i^\alpha$ $(1 \leqq i \leqq n, 1 \leqq \alpha \leqq m)$. Furthermore, the first m coordinates in U can be considered as normal coordinates in U', and the last $n - m$ coordinates are local coordinates for the fibres.*

Proof. Let $(e_i)_{1 \leqq i \leqq n}$ be an orthonormal base for $M(P)$ such that the first m vectors span $M_H(P)$, and the last $n - m$ vectors span $M_V(P)$. Then $f_*(P)e_i = e'_i$ $(1 \leqq i \leqq m)$ form an orthonormal base for $M'(f(P))$, and we can construct the associated normal coordinates in some neighborhood U'. According to Hermann [13] the unique horizontal lift to P of any geodesic of M' starting at $f(P)$ is a geodesic of M; one determined by e'_i lifts to one determined by e_i. We now use the local product structure to define a coordinate chart U in which the fibres have the desired property. Note that (unless the fibres are totally geodesic) we cannot generally require that the coordinates in U be normal.

PROPOSITION. *Let $f\colon M \to M'$ be a Riemannian fibration. Then* $e(f)$

$-n/2$. If for any $P \in M$ we let F_P denote the fibre through P and $i_P \colon F_P \to M$ the inclusion map, then

$$\tau(f)(Q) = -f_*(Q)\tau(i_P)(Q) \text{ for all } Q \in F_P.$$

In particular, f is harmonic if and only if all fibres are minimal submanifolds.

Proof. In § 2D we have seen that the tension field of any Riemannian immersion is perpendicular to the submanifold. Thus for every $Q \in F_P$ we have $\tau(i_P)(Q) \in M_H(Q)$. If we use a split coordinate system as in the previous lemma, we see that

$$f_*\tau(i_P)(P)^\gamma = \sum_{k=1}^{m} \delta_k{}^\gamma \tau(i_P)^k(P).$$

The proposition follows by direct calculation.

Examples. All covering maps are harmonic; in particular, the identity map is harmonic. which amounts to saying that Cartan's *vitesse* is a harmonic 1-form. If $\mathfrak{F}^r(M) \to M$ is the bundle of orthonormal r-frames of M, then it is known (Lichnerowicz [15]) that with its natural Riemannian structure on $\mathfrak{F}^r(M)$ the fibres are minimal, whence that fibre map is harmonic. Vector bundle maps are harmonic. Every homogeneous Riemannian fibration is harmonic, for the fibres are always totally geodesic, and therefore minimal.

(D) *Maps into flat manifolds.* Let us take for M' the unit circle S^1; *we will construct harmonic representatives in every homotopy class of maps $M \to S'$.* First of all, it is well known that the set $[M, S^1]$ of these homotopy classes forms an abelian group canonically isomorphic to the first integral cohomology group $H^1(M)$. Secondly, every such cohomology class is canonically represented by a harmonic 1-form on M.

Suppose M is connected, and fix a point $P_0 \in M$. Given any such harmonic 1-form ω on M and any smooth path γ_P from P_0 to a point $P \in M$, we define the number

$$f(P) = \int_{\gamma_P} \omega.$$

A different choice $\bar{\gamma}_P$ of γ_P may give a different number $\bar{f}(P)$, but

$$\bar{f}(P) - f(P) = \int_{\bar{\gamma}_P - \gamma_P} \omega$$

is an integer since the periods of ω are integral; consequently, ω determines a well defined map $f_\omega \colon M \to S^1$ by letting $f_\omega(P)$ be the residue class modulo 1 of $f(P)$.

Now since f_ω is harmonic, every $P \in M$ has a neighborhood U in which $df = \omega$; thus $\Delta f = \delta df + d\delta f = \delta\omega = 0$ in U. It is easy to see that $\omega \rightarrow f_\omega$ establishes an isomorphism $[M, S^1] = H^1(M)$.

To define *harmonic representatives of the homotopy classes of maps* $M \rightarrow T^m$, *the flat m-torus,* we merely take m-fold products of harmonic maps $M \rightarrow S^1$, using § 5C below; the existence of these harmonic representatives was first proved by F. B. Fuller [10]. More generally, any compact flat manifold M' is covered by T^m, by a theorem of Bieberbach, and any homotopy class of maps $M \rightarrow M'$ which can be lifted to maps $M \rightarrow T^m$ has harmonic representatives obtained by composition with the projection $T^m \rightarrow M'$.

If M and M' are both flat, then the only harmonic maps $M \rightarrow M'$ are those which are locally linear, as can be seen from the maximum principle.

(E) *Maps of Euclidean spheres.* Let $S^n(r)$ denote the Euclidean n-sphere of radius r; write $S^n = S^n(1)$. Then the homotopy classes of maps of S^n into itself are classified by their degrees. We consider now the problem of constructing explicitly harmonic maps of a given degree k; it is a simple matter to modify the following remarks to include the case of maps $S^n(r) \rightarrow S^n(r')$.

If $(x^1, \cdots, x^{n+1})$ are Euclidean coordinates and $\eta\colon S^n(r) \rightarrow S^n(r')$ is a map given by $\eta(x^1, \cdots, x^{n+1}) = (x^1 r'/r, \cdots, x^{n+1} r'/r)$, then

$$e(\eta) = n/2(r'/r)^2.$$

We will henceforth refer points of S^n to coordinates (θ, ϕ), where θ denotes colatitude $(0 \leqq \theta \leqq \pi)$, and ϕ a point on the equator S^{n-1} of S^n. Furthermore, corresponding to the integer k let $\Phi\colon S^{n-1}(r) \rightarrow S^{n-1}(r')$ be the $(n-2)$-fold suspension of the map $S^1(r) \rightarrow S^1(r')$ defined by $\phi \rightarrow (\cos k\phi, \sin k\phi)$; since the degree is invariant under suspension, Φ has degree k. An elementary calculation gives

$$e(\Phi) = \frac{(k^2 + n - 2)}{2}(r'/r)^2. \tag{17}$$

We are interested in maps $f\colon S^n \rightarrow S^n$ of the form $(\theta, \phi) \rightarrow (\Theta, \Phi)$, where Θ is a function of θ alone, and Φ is defined for $r = r' = 1$. From (17) we see that on $S^{n-1}(\sin\theta)$ we have $2e(f) = (k^2 + n - 2)\sin^2\Theta/\sin^2\theta$, whence

$$\begin{aligned} E(f) &= \tfrac{1}{2}\int_0^\pi \int_{S^{n-1}(\sin\theta)} \left[\left(\frac{d\Theta}{d\theta}\right)^2 + (k^2 + n - 2)\frac{\sin^2\Theta}{\sin^2\theta}\right] ds^{n-1} d\theta. \\ &= \frac{V(S^{n-1})}{2}\int_0^\pi \sin^{n-1}\theta \left[\left(\frac{d\Theta}{d\theta}\right)^2 + (k^2 + n - 2)\frac{\sin^2\Theta}{\sin^2\theta}\right] d\theta. \end{aligned} \tag{18}$$

In case $n = 1$ the existence and properties of harmonic maps of degree k is elementary.

We now consider the case $n = 2$. We have $g_{11} = 1$, $g_{12} = 0 = g_{21}$, $g_{22} = \sin^2\theta$. The tension field of any map f is given by its components

$$\tau(\Theta) = \Delta\Theta - \sin\Theta\cos\Theta\left[\frac{1}{\sin^2\theta}\left(\frac{\partial\Phi}{\partial\phi}\right)^2 + \left(\frac{\partial\Phi}{\partial\theta}\right)^2\right],$$

$$\tau(\Phi) = \Delta\Phi + 2\cot\Theta\left[\frac{1}{\sin^2\theta}\frac{\partial\Phi}{\partial\phi}\frac{\partial\Theta}{\partial\phi} + \frac{\partial\Phi}{\partial\theta}\frac{\partial\Theta}{\partial\theta}\right].$$

For the special maps under consideration we find that for a harmonic map f $\tau(\Phi) = \Delta\Phi = d^2\Phi/d\phi^2 = 0$, which means that the choice $\Phi = k\phi$ is compatible with our above selection. Substituting in the first equation we find

$$\tau(\Theta) = \frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) - k^2\frac{\sin\Theta\cos\Theta}{\sin^2\theta};$$

the only solution of $\tau(\Theta) = 0$ regular at the poles is

$$\Theta = 2\arctan\left[c(\tan\theta/2)^{\pm k}\right] \text{ with } c > 0. \tag{19}$$

Then $d\Theta/d\theta = \pm k\sin\Theta/\sin\theta$, from which we can conclude that $E(f) = 4\pi\,|\,k\,|$; note that it is independent of c. Finally, the integral formula for the degree of a map shows that degree $f = \pm k$. We observe that for $k = 0$ and all $c > 0$ the map is constant; for $k = \pm 1$ and $c = 1$ the map is the identity and the antipodal map, respectively.

Remark. Although the above construction does exhibit a harmonic map in every homotopy class, it does not begin to exhaust their topological interest. For instance, with the uniform topology on $\mathcal{H}(S^2, S^2)$, the component $\mathcal{H}_0(S^2, S^2)$ of those maps of degree 0 has infinite cyclic homology:

$$H_1(\mathcal{H}_0(S^2, S^2)) = \pi_1(\mathcal{H}_0(S^2, S^2)) = \pi_3(S^2) = Z,$$

generated by the Hopf map. A harmonic representative of that generator should have positive Morse index.

Consider now the case $n \geqq 3$. We do not know how to construct harmonic maps of degree $\geqq 2$. Incidentally, the suspensions of the above maps (and their "compressed" suspensions below) are generally not harmonic. We propose now to show that *the functional* $E\colon \mathcal{H}(S^n, S^n) \to R$ *does not have an absolute minimum on the component* $\mathcal{H}_k(S^n, S^n)$ *for* $k \neq 0$ *and* $n \geqq 3$; that this phenonmenon could be illustrated explicitly and simply by maps of spheres was suggested to us by C. B. Morrey.

Let $f_c(\theta, \phi) = (\Theta_c(\theta), \Phi)$ be the map of $S^n \to S^n$ defined using (19) with exponent $k \neq 0$. Thus for small $c > 0$ the map f_c compresses most of S^n into a small cell centered at the pole, and that compression takes place along longitudes.

The energy integral (18) now takes the form

$$E(f_c) = \frac{V(S^{n-1})(2k^2+n-2)}{2} \int_0^\pi \sin^{n-1}\theta \cdot \left(\frac{\sin\Theta_c(\theta)}{\sin\theta}\right)^2 d\theta.$$

For $n \geqq 3$ we have

$$E(f_c) \leqq \frac{V(S^{n-1})(2k^2+n-2)}{2} \int_0^\pi \sin^2\Theta_c(\theta)\, d\theta.$$

The following lemma shows that $\lim_{c\to 0} E(f_c) = 0$; i.e., for $n \geqq 3$ there are maps in $\mathcal{H}_k(S^n, S^n)$ of arbitrarily small energy. But for $k \neq 0$ there is no map $f \in \mathcal{H}_k(S^n, S^n)$ with zero energy, for such an f is constant, and therefore has degree 0.

LEMMA. $\int_0^\pi \sin^2\Theta_c(\theta)\, d\theta \to 0$ *as* $c \to 0$.

Proof. For any $\epsilon > 0$ let $\rho = \pi - \epsilon/2$. There is a number K such that $0 \leqq (\tan\theta/2) \leqq K$ for all $0 \leqq \theta \leqq \rho$. It follows that $\Theta_c(\theta) \leqq 2\arctan(cK)$, whence there is a number $c_\epsilon > 0$ for which $0 \leqq \sin^2\Theta_c(\theta) < \epsilon/2\rho$ if $0 < c \leqq c_\epsilon$. Thus

$$\int_0^\rho \sin^2\Theta_c(\theta)\, d\theta + \int_\rho^\pi \sin^2\Theta_c(\theta)\, d\theta \leqq \int_0^\rho \epsilon/2\rho\, d\theta + \int_\rho^\pi 1\, d\theta \leqq \epsilon.$$

5. The composition of maps.

(A) The following computation is elementary.

LEMMA. *If* $f\colon M \to M'$ *and* $f'\colon M' \to M''$ *are any smooth maps, then their fundamental forms satisfy*

$$(f'\circ f)_{;ij}{}^{a} = f_{;ij}{}^{\gamma} f'_{\gamma}{}^{a} + f'_{;\alpha\beta}{}^{a} f_i{}^{\alpha} f_j{}^{\beta}. \tag{20}$$

COROLLARY. *The composition of totally geodesic maps is totally geodesic. The inverse of a totally geodesic diffeomorphism is totally geodesic.*

COROLLARY.

$$\tau(f'\circ f)^{a} = \tau(f)^{\gamma} f'_{\gamma}{}^{a} + g^{ij} f'_{;\alpha\beta}{}^{a} f_i{}^{\alpha} f_j{}^{\beta}. \tag{21}$$

If f *is harmonic and* f' *totally geodesic, then* $f'\circ f$ *is harmonic.*

In general, however, we do not expect the composition of harmonic maps to be harmonic, as the following example shows:

Example. Let T^2 be the flat 2-torus parametrized by the angles (θ, ϕ) with $0 \leqq \theta, \phi, < 2\pi$. Let $f' \colon T^2 \to S^3$ be defined by

$$f'(\theta, \phi) = (\cos\theta, \sin\theta, \cos\phi, \sin\phi)/\sqrt{2},$$

considered as a point in R^4. Then *f' defines a Riemannian imbedding of T^2 in S^3, which is a minimal but not a totally geodesic imbedding.* To see that f' is harmonic we show that $\tau(f')$ is perpendicular to $S^3(f'(P))$ in R^4, and then appeal to Proposition 5B below. Namely, because T^2 is flat we have

$$\tau(f')^a = \frac{\partial^2 f'^a}{\partial\theta^2} + \frac{\partial^2 f'^a}{\partial\phi^2} = -f'^a \qquad (1 \leqq a \leqq 4)$$

whence $\tau(f')(P)$ is directed along the radius of S^3 at $f(P)$. On the other hand, T^2 is not totally geodesic in S^3, for the map $f \colon S^1 \to T^2$ defined by $f(0) = (\theta, 0)$ is a geodesic of T^2; it does not lie in any 2-plane through 0 in R^4, and is therefore not a geodesic of S^3. In particular, *$f \colon S^1 \to T^2$ and $f' \colon T^2 \to S^3$ are both harmonic maps, and their composition $f' \circ f$ is not.*

(B) PROPOSITION. *If $f' \colon M' \to M''$ is a Riemannian immersion, then for any map $f \colon M \to M'$ we have $E(f) = E(f' \circ f)$. The tension field $\tau(f)$ is the projection on M' of the tension field $\tau(f' \circ f)$.*

Proof. The first statement follows from the equation

$$e(f' \circ f) = \tfrac{1}{2}\langle g, f^* \circ f'^* g'' \rangle = e(f).$$

The second statement is a consequence of (21); for if f' is a Riemannian immersion, then the right member is the decomposition of $\tau(f' \circ f)$ into horizontal and vertical components because $(f'_{;\alpha\beta}{}^{a})$ is the second fundamental form of f'.

COROLLARY. *A map $f \colon M \to M'$ is harmonic if and only if $\tau(f' \circ f)$ is perpendicular to $M'(f(P))$ for all $P \in M$.*

This generalizes the classical fact that a curve in M' is a geodesic if and only if its curvature vector in M'' is always perpendicular to M'.

(C) PROPOSITION. *Let $f' \colon M' \to M''$ be a Riemannian fibration with totally geodesic fibres. Then for any map $f \colon M \to M'$ we have*

$$\tau(f' \circ f) = f'_{*}(\tau(f)).$$

This is immediate, because we can in the present situation take split normal coordinates in Lemma 4C.

COROLLARY. *If* $M = M''$ *and* $f\colon M \to M'$ *is a section, then* $f' \circ f = 1$ *is harmonic, and therefore* $\tau(f)$ *is always vertical.*

Example. If we view a smooth r-form ω on M' as a section of the bundle $\mathcal{J}^{[r]}(M')$ of r-covectors of M', then the condition that ω be harmonic in the sense of de Rham-Hodge is generally different from the condition $\tau(\omega) = 0$, using (say) the Riemannian structure on $\mathcal{J}^{[r]}(M')$ of Sasaki [25]. However, these two concepts do coincide if M' is flat.

Example. A map $f\colon M \to M' \times M''$ into a Riemannian product has a canonical decomposition $f(P) = (f'(P), f''(P))$ for all $P \in M$. Then f is harmonic if and only if both components f', f'' are harmonic. For instance, seee Proposition 2E.

(D) Let us suppose that M' is a Riemannian submanifold of M'' and that the imbedding is proper; i. e., such that the inverse image of any compact subset of M'' is compact in M'. Since M' is complete, there is a positive smooth function $\rho\colon M' \to R$ such that for any $P' \in M'$ the set

$$\{P'' \in M'' : r'(P', P'') \leqq \rho(P')\}$$

is geodesically convex in M''; if M' is compact, then of course we can suppose that ρ is a positive constant. For each $P' \in M'$ let $D_{P'}$ denote the closed ball of dimension $q - m$ ($q = \dim M''$) consisting of all geodesic segments of length $\leqq \rho(P')$ emanating from P' and perpendicular to $M'(P')$. The following result is well known and elementary.

LEMMA. *Let* $N = \cup \{D_{P'} : P' \in M'\}$; *then* N *is a neighborhood of* M' *in* M'', *and the obvious map* $\pi\colon N \to M'$ *defines a smooth fibre bundle over* M' *whose structure group is* O_{q-m} *and whose fibres are closed balls.*

Taking into account Proposition 5C we have the

PROPOSITION. *Let* $f'\colon M' \to M''$ *be a proper Riemannian imbedding and* $\pi\colon N \to M'$ *a normal tubular neighborhood. Then for any map* $f\colon M \to N$ *the composition* $\pi \circ f$ *is harmonic if and only if* $\tau(f)$ *is vertical.*

Chapter II. Deformations of Maps.

6. Deformations by the heat equation.

(A) This chapter is devoted to *the fundamental problem of deforming a given map into a harmonic map*; i.e., into a smooth map $f\colon M \to M'$ satisfying the nonlinear elliptic equation

$$(1) \qquad \tau(f) = 0.$$

We begin by discussing general methods of attack.

The interpretation given in § 2A of the tension field $\tau(f)$ as the contravariant representative of the differential of E at f suggests that we try to invent gradient lines of E in a suitable function space of maps from M to M', and then to prove that these trajectories lead to critical points of E. We propose the following method for realizing such an attempt, which we now outline briefly. We do not pursue this method in the present paper, although the qualitative results are essentially those of the following sections.

Let $\mathfrak{H}^r(M,M')$ denote the function space of all maps from M to M' whose partial derivatives (relative to fixed coordinate coverings) of orders $\leq r$ are square integrable. An inequality of Sobolev insures that if $2r > \dim M$ then the maps in $\mathfrak{H}^r(M,M')$ are continuous, and its topology is larger than the uniform topology. It can be shown that the space $\mathfrak{H}^r(M,M')$ admits an infinite dimensional Riemannian manifold structure modeled on a separable Hilbert space, and that $E\colon \mathfrak{H}^r(M,M') \to R$ is a differentiable function. If τ^r is its gradient field on $\mathfrak{H}^r(M,M')$; i.e., $\nabla_v E(f) = -\langle \tau^r(f), v\rangle$ for all vectors v in the tangent space at f, then the ordinary differential equation

$$\frac{df_t}{dt} = \tau^r(f_t)$$

has a local solution which is unique; furthermore, $E(f_t)$ is a decreasing function of t. Under suitable curvature restrictions the solutions are globally defined. If each trajectory f_t is relatively compact, then it has a limit point a harmonic map. Thus these trajectories define a canonical homotopy of the initial map onto a harmonic map; moreover, such trajectories enjoy the 1-parameter group property. We observe that the critical points of E are just the zeros of all τ^r (for any $r > \dim M/2$).

Now the function space $\mathfrak{H}^0(M,M')$ is *not* a manifold, although with every map f we have the Hilbert space $\mathfrak{H}^0(f)$ of vector fields along f defined in § 2A. In particular, $\tau^0(f) = \tau(f)$ is in $\mathfrak{H}^0(f)$. In analogy with the above outline we are led to consider the nonlinear parabolic equation

$$(2) \qquad \frac{\partial f_t}{\partial t} = \tau(f_t) \qquad (t_0 < t < t_1).$$

The study of this equation is our primary object in the following sections. We will find that the properties of the trajectories of (2) include most of those mentioned as belonging to the trajectories of τ^r $(2r > \dim M)$. (There is one basic difference: The solutions of (2) are generally defined only for non-negative time, whereas the trajectories of τ^r are always defined for an open time interval around $t=0$.)

Remark. In the calculus of variations a standard method (Morrey [19]) of establishing the existence of a minimum of E for a given class of maps is to take the space $\mathfrak{H}^1(M,M')$ and to introduce on it a weak topology relative to which 1) E is lower semi-continuous, and 2) there are sufficiently many compact sets. That approach works well for $\dim M=1$ or 2; however, the example given in § 4E shows that it will not work in general for $n \geqq 3$.

(B) PROPOSITION. *If* $(t,P)\rightarrow f_t(P)$ *is a map of* $(t_0,t_1)\times M\rightarrow M'$ *which is* C^1 *on the product manifold and* C^2 *on* M *for each* t, *and if that map satisfies* (2), *then it is* C^∞.

We will refer to such an f_t as a *solution* of (2).

This follows from Friedman [9, Th. 4 and 5] provided the second derivatives $f_{ij}{}^\alpha$ of the f^α are Hölder-continuous. But we can represent the local functions f^α by Green's formula, using the fundamental solution of the heat equation $\Delta u-\partial u/\partial t=0$ as in §§ 9-10 below. The required Hölder-continuity is then established by standard techniques from the properties of the potentials involved (Pogorzelski [24], Dressel [7], Gevrey [11, No. 8]).

(C) Let $f_t\colon M\rightarrow M'$ satisfy (2); the subscript t refers to the deformation parameter (we will always indicate explicitly differentiation with respect to t). Then from (4) of § 2B we have

$$\frac{dE(f_t)}{dt}=-\int_M\langle\tau(f_t),\frac{\partial f_t}{\partial t}\rangle*1=-\int_M\left|\frac{\partial f_t}{\partial t}\right|^2*1.$$

If D/dt denotes covariant differentiation along paths in M', then for each $P\in M$ the curvature vector of the path $t\rightarrow f_t(P)$ is given by $D(\partial f_t/\partial t)/dt$.

PROPOSITION. *If* $f_t\colon M\rightarrow M'$ *satisfies* (2), *then the energy* $E(f_t)$ *is a strictly decreasing function; i. e.,* $dE(f_t)/dt<0$ *except for those values of* t *for which* $\tau(f_t)=0$. *Furthermore, its second derivative expresses the average angle between the tension field and the curvature vectors of the deformation paths:*

$$\frac{d^2E(f_t)}{dt^2}=-2\int_M\langle\frac{D}{dt}\left(\frac{\partial f_t}{\partial t}\right),\tau(f_t)\rangle*1.$$

LEMMA. *Let* $f_t\colon M\rightarrow M'$ *be an arbitrary deformation for* $t\in(t_0,t_1)$. *If we let*

$$\frac{D}{\partial x^i}\left(\frac{\partial f_t}{\partial t}\right)^\alpha=\frac{\partial^2 f_t{}^\alpha}{\partial x^i\partial t}+\Gamma'_{\mu\nu}{}^\alpha\frac{\partial f_t{}^\mu}{\partial x^i}\frac{\partial f_t{}^\nu}{\partial t},$$

then

$$\frac{d^2E(f_t)}{dt^2} = \int_M g^{ij} < \frac{D}{\partial x^j}\left(\frac{\partial f_t}{\partial t}\right), \frac{D}{\partial x^j}\left(\frac{\partial f_t}{\partial t}\right) > *1$$

$$- \int_M g^{ij} R'_{\alpha\mu\beta\nu} f_{ti}{}^{\alpha} f_{tj}{}^{\beta} \frac{\partial f_t{}^{\mu}}{\partial t} \frac{\partial f_t{}^{\nu}}{\partial t} *1 - \int_M < \frac{D}{\partial t}\left(\frac{\partial f_t}{\partial t}\right), \tau(f_t) > *1.$$

This follows from a direct calculation of $\partial^2 e(f_t)/\partial t^2$ and an application of Green's divergence theorem.

Because the first integral of the right member is always non-negative, we have on appeal to the above proposition the

THEOREM. *If $f_t \colon M \to M'$ satisfies the heat equation (2), then*

$$(3) \quad \begin{aligned} \frac{1}{2}\frac{d^2E(f_t)}{dt^2} &= \int_M g^{ij} < \frac{D}{\partial x^i}\left(\frac{\partial f_t}{\partial t}\right), \frac{D}{\partial x^j}\left(\frac{\partial f_t}{\partial t}\right) > *1 \\ &- \int_M g^{ij} R'_{\alpha\mu\beta\nu} f_{ti}{}^{\alpha} f_{tj}{}^{\beta} \frac{\partial f_t{}^{\mu}}{\partial t} \frac{\partial f_t{}^{\nu}}{\partial t} *1. \end{aligned}$$

In particular, if M' has non-positive sectional curvature, then $d^2E(f_t)/dt^2 \geqq 0$. If t is a value for which equality holds, then $\tau(f_t)$ is a covariant constant; i. e.,

$$\frac{D}{\partial x^i}\left(\frac{\partial f_t(P)}{\partial t}\right) = 0 \text{ for all } P \in M \text{ and } (1 \leqq i \leqq n).$$

COROLLARY. *If M' has non-positive sectional curvature and if f_t satisfies (2) for all $t \geqq t_0$, then*

$$\frac{dE(f_t)}{dt} \to 0 \text{ as } t \to \infty.$$

(D) We have seen in § 5 that in supposing M' contained in a larger manifold M'' we do not alter the energy of a map $f \colon M \to M'$. That suggests that we still have control of the energy and the tension of deformations of f which take place in a normal tubular neighborhood N of M' in M''.

If $f' \colon M' \to M''$ denotes the imbedding, then the induced tangent vector bundle is $f'^{-1}\mathcal{T}(M'') = \mathcal{T}(M') \oplus \mathcal{N}(M'', M')$, where the second summand is the normal bundle of M' in M''. Then $\mathcal{N} \to M'$ is Riemannian-connected in the sense of § 3A. There is a canonical vector field $\rho \colon N \to \mathcal{N}$ covering the projection map $\pi \colon N \to M'$ defined by assigning to each $Q'' \in N$ the unique vector $\rho(Q'') \in M''(\pi(Q''))$ such that $\exp_{\pi(Q'')}(\rho(Q'')) = Q''$.

Suppose now that $f_t \colon M \to N$ is a smooth deformation $(t_0 \leqq t < t_1)$. Then $\rho \circ f_t$ is a vector field on the map $\pi \circ f_t$, and using the harmonic integral theory of § 3 applied to the vector bundle $\mathfrak{N} \to M'$, we can define its Laplacean $\bar{\Delta}(\rho \circ f_t)$. Let

$$\bar{L}(\rho \circ f_t) = \bar{\Delta}(\rho \circ f_t) - \frac{\bar{D}}{\partial t}(\rho \circ f_t),$$

where $\bar{D}/\partial t$ is the covariant derivative in N along the path $\pi \circ f_t$.

We now establish the following stability property of deformations.

THEOREM. *Let N be a normal tubular neighborhood of M' in M'', and suppose that $f_t\colon M \to N$ is a smooth deformation $(t_0 \leqq t < t_1)$. If $\bar{L}(\rho \circ f_t)$ is always horizontal in $\mathcal{J}(M'')$ and if $f_{t_0}\colon M \to M'$, then $f_t\colon M \to M'$ for all $t_0 \leqq t < t_1$.*

Proof. We apply Green's theorem to $u, v \in A^0(M, (\pi \circ f_t)^{-1}\mathfrak{N})$:

$$\int_M \langle u, \bar{\Delta} v\rangle *1 = -\int_M \langle \bar{d}u, \bar{d}v\rangle *1.$$

The hypotheses imply that $\langle \rho \circ f_t, \bar{L}(\rho \circ f_t)\rangle = 0$ for all t, so that

$$0 = \int_M \langle \rho \circ f_t, \bar{\Delta}(\rho \circ f_t)\rangle *1 - \int_M \langle \rho \circ f_t, \frac{\bar{D}}{\partial t}(\rho \circ f_t)\rangle *1.$$

Therefore

$$\frac{1}{2}\frac{d}{dt}\int_M \langle \rho \circ f_t, \rho \circ f_t\rangle *1 = \int_M \langle \rho \circ f_t, \bar{\Delta}(\rho \circ f_t)\rangle *1$$
$$= -\int_M |\bar{d}(\rho \circ f_t)|^2 *1 \leqq 0.$$

I.e., $\int_M |\rho \circ f_t|^2 *1$ is a non-negative, non-increasing function of t, and it is zero for $t = t_0$. We conclude that $\rho \circ f_t = 0$ for all $t_0 \leqq t < t_1$, which states precisely that every f_t maps M into M'.

7. Global equations.

(A) We now occupy ourselves with the problem of replacing equations (1) and (2), which in terms of local coordinates on M and M' are local systems of equations, by some much more tractable global systems.

Remark. Assume for the moment that we have an isometric imbedding $w\colon M' \to \boldsymbol{R}^q$ for some q, which we can always do by a theorem of Nash [21]. Then as in Proposition 5B we find that equation (2) is satisfied for a deformation $f_t\colon M \to M'$ when and only when the composition $W_t = w \circ f_t$ satisfies the condition that the vectors $LW_t = \Delta W_t - \frac{\partial W_t}{\partial t}$ are perpendicular to M'; see

Lemma 7B below. When expressed in terms of the coordinates of $\boldsymbol{R}^q$, that condition gives rise to a global parabolic system of equations of the type $(\tilde{2})$. On the other hand, the assumption of an isometric imbedding apparently affords no real simplification in our theory, and we will not make it. We proceed with an elementary imbedding convenient for our purposes.

Suppose that M' is smoothly and properly imbedded in some Euclidean space $\boldsymbol{R}^q$ by a map $w\colon M' \to \boldsymbol{R}^q$.

LEMMA. *Given such an imbedding of M' in $\boldsymbol{R}^q$, it is always possible to construct a smooth Riemannian metric $g'' = (g''_{ab})_{1\leq a,b\leq q}$ on a tubular neighborhood N of M' so that N is Riemannian fibred.*

Proof. Let N be any tubular neighborhood of M' constructed using the Euclidean structure of $\boldsymbol{R}^q$; let $\pi\colon N \to M'$ be the projection map. It suffices to construct an appropriate smooth inner product in each space $\boldsymbol{R}^q(P')$ for all $P' \in M'$, for we can translate that tangent space to any point $Q' \in N$ along the straight line segment (necessarily contained in N) from $P' = \pi(Q')$ to Q'. We take g' in $M'(P')$ and the induced Euclidean metric in the (Euclidean) orthogonal complement of $M'(P')$ in $\boldsymbol{R}^q(P')$, and take their sum in $\boldsymbol{R}^q(P')$.

In terms of local coordinates $(y^1,\cdots,y^m)$ on M' that metric can be described as follows: Write $w(P') = (w^1(P'),\cdots,w^q(P'))$; then

$$\sum_{c=1}^{q} \frac{\partial w^c}{\partial y^\alpha}\,\frac{\partial w^c}{\partial y^\beta}$$

is the metric on M' induced from the Euclidean metric. Let $t_a{}^\alpha$ be the unique solution of

$$\sum_{c,\alpha} \frac{\partial w^c}{\partial y^\alpha}\,\frac{\partial w^c}{\partial y^\beta}\,t_a{}^\alpha = \frac{\partial w^a}{\partial y^\beta}.$$

We have the duality relation $\dfrac{\partial w^a}{\partial y^\alpha} t_a{}^\beta = \delta_\alpha{}^\beta$, and the metric tensor is

$$g''_{ab} = \delta_{ab} - \sum_{c=1}^{q} \frac{\partial w^c}{\partial y^\alpha}\,\frac{\partial w^c}{\partial y^\beta} t_a{}^\alpha t_b{}^\beta + g'_{\alpha\beta} t_a{}^\alpha t_b{}^\beta.$$

Then

$$g''_{ab}\frac{\partial w^a}{\partial y^\alpha}\,\frac{\partial w^b}{\partial y^\beta} = g'_{\alpha\beta},$$

so that g'' does induce g' on M'. For any vector $v \in \boldsymbol{R}^q(P')$ satisfying $\sum_{a=1}^{q} v^a \dfrac{\partial w^a}{\partial y^\alpha} = 0$ $(1 \leqq \alpha \leqq m)$ we have $\sum_{a=1}^{q} v^a t_a{}^\alpha = 0$ also, so that

$$g''_{ab} v^a \frac{\partial w^b}{\partial y^\alpha} = 0 \text{ and } g''_{ab} v^a v^b = \sum_a (v^a)^2.$$

Thus g'' has the desired properties.

(B) LEMMA. *Let* $f_t\colon M \to M'$ *be any smooth deformation, and let* $W_t\colon M \to N$ *be the composition* $w \circ f_t = W_t$. *Then* $\tau(f_t) = \partial f_t/\partial t$ *if and only if* $\tau(W_t) - \partial W_t/\partial t$ *is perpendicular to* $M'(f_t(P))$ *for all* $P \in M$.

Proof. The argument is essentially that of Proposition 5B; we choose local coordinates on M' and obtain

$$g'_{\gamma\lambda}\Big(\tau(f_t)^\lambda - \frac{\partial f_t{}^\lambda}{\partial t}\Big) = g''_{cd}\frac{\partial w^c}{\partial y^\gamma}\,\tau(W_t)^d - g'_{\gamma\lambda}\frac{\partial f_t{}^\lambda}{\partial t}$$

$$= g''_{cd}\frac{\partial w^c}{\partial y^\gamma}\,\Big(\tau(W_t)^d - \frac{\partial W^d}{\partial t}\Big).$$

In terms of the coordinates of R^q the differential of the projection map $\pi\colon N \to M'$ has components $\pi_a{}^c = \partial\pi^c/\partial w^a$. Its covariant differential in terms of the metric g'' is

$$\pi_{ab}{}^c = \frac{\partial^2\pi^c}{\partial w^a \partial w^b} - \Gamma''_{ab}{}^d\pi_d{}^c \qquad (1 \leqq a, b, c \leqq q).$$

The map $\rho\colon N \to \boldsymbol{R}^q$ defined (as in § 6D) by $\rho(Q') = Q' - \pi(Q')$ assigns to each $Q' \in N$ a vector perpendicular to $M'(\pi(Q'))$, and $\rho_a{}^c + \pi_a{}^c = \delta_a{}^c$,

$$\rho_{ab}{}^c + \pi_{ab}{}^c + \Gamma''_{ab}{}^c = 0.$$

For its restriction to M' we obtain

$$-\rho_{ab}{}^c w^a w^b = \frac{\partial^2\pi^c}{\partial w^a \partial w^b}\; w_\alpha{}^a w_\beta{}^b = \rho_b{}^c\,\frac{\partial^2\pi^b}{\partial y^\alpha \partial y^\beta}\,,$$

where $w_\alpha{}^a = \partial w^a/\partial y^\alpha$. For each α, β the right member defines a vector perpendicular to M'; because the $w_\alpha{}^a$ span the tangent space to M' in which they lie, we obtain the

LEMMA. *For any vectors* $u, v \in M'(Q')$ *the vector whose components are* $\rho_{ab}{}^c u^a v^b$ *is perpendicular to* $M'(Q')$.

LEMMA. *For any map* $W_t\colon M \to N$ *with image in* M' *let* ξ *be the vector with components*

$$\xi^c = LW_t{}^c - \pi_{ab}{}^c W_{ti}{}^a W_{tj}{}^b g^{ij},$$

where $L = \Delta - \partial/\partial t$ *is the heat operator on* M. *Then* ξ *is tangent to* M'.

Proof. As in Proposition 5B we have

$$L(\rho(W_t))^d = \rho_c{}^d L(W_t)^c + \frac{\partial^2\rho^d}{\partial w^a \partial w^b}\, W_{ti}{}^a W_{tj}{}^b g^{ij}$$

(5)
$$= \rho_c{}^d L(W_t)^c - \frac{\partial^2 \pi^d}{\partial w^a \partial w^b} W_{ti}{}^a W_{tj}{}^b g^{ij}$$

But $\rho(W_t) \equiv 0$, whence the left member vanishes. Since the second term in the right member is a normal vector, we have

$$\rho_c{}^d L(W_t)^c = \rho_e{}^d \frac{\partial^2 \pi^e}{\partial w^a \partial w^b} W_{ti}{}^a W_{tj}{}^b g^{ij}.$$

It follows that

$$\rho_c{}^d \xi^c = \left(\rho_e{}^d \frac{\partial^2 \pi^e}{\partial w^a \partial w^b} - \rho_e{}^d \pi_{ab}{}^e\right) W_{ti}{}^a W_{tj}{}^b g^{ij} = 0$$

for $1 \leqq d \leqq q$. Thus ξ has normal component 0.

PROPOSITION. *A map $f\colon M \to M'$ satisfies* (1) *if and only if the composition $W = w \circ f$ satisfies*

$(\tilde{1})$
$$\Delta W^c = \pi_{ab}{}^c W_i{}^a W_j{}^b g^{ij}$$

in terms of local coordinates on M.

A deformation $f_t\colon M \to M'$ $(t_0 < t < t_1)$ satisfies (2) *if and only if $W_t = w \circ f_t$ satisfies*

$(\tilde{2})$
$$L(W_t)^c = \pi_{ab}{}^c W_{ti}{}^a W_{tj}{}^b g^{ij} \qquad (t_0 < t < t_1).$$

Proof. It suffices to establish the equivalence of (2) and $(\tilde{2})$. For that we take any deformation f_t and compute

$$\begin{aligned}\left(\tau(W_t)^c - \frac{\partial W_t{}^c}{\partial t}\right) g''{}_{cd} w_\gamma{}^d &= \left(L(W_t)^c + \Gamma''{}_{ab}{}^c W_{ti}{}^a W_{tj}{}^b g^{ij}\right) g''{}_{cd} w_\gamma{}^d \\ &= \left(L(W_t)^c - \pi_{ab}{}^c W_{ti}{}^a W_{tj}{}^a g^{ij}\right) g''{}_{cd} w_\gamma{}^d,\end{aligned}$$

by the second lemma. If W_t satisfies $(\tilde{2})$, then $\tau(W_t)(P) - \partial W_t(P)/\partial t$ is perpendicular to $M'(f_t(P))$, whence by the first lemma f_t satisfies (2). Conversely, if f_t satisfies (2) and we define ξ as in the third lemma, then our equation shows that $\xi(P)$ is perpendicular to $M'(f_t(P))$, whence it must be 0.

(C) The following result is an application of Theorem 6D. In order to have a proof avoiding the use of vector-bundle-valued harmonic forms, we can start with (5) and substitute $(\tilde{2})$ to obtain

$$L(\rho(W_t))^c = -\pi_d{}^c \left(\frac{\partial^2 \pi^d}{\partial w^a \partial w^b} + \rho_c{}^d \Gamma''{}_{ab}{}^e\right) W_{ti}{}^a W_{tj}{}^b g^{ij},$$

using the projection relation $\rho_d{}^c \pi_e{}^d = \pi_d{}^c \rho_e{}^d$. It follows that $\rho^c \cdot L(\rho(W_t))^c = 0$, so that

$$\sum_{c} \rho^{c} \cdot \Delta(\rho(W_t))^{c} = \tfrac{1}{2}\frac{\partial}{\partial t} \sum_{c} (\rho(W_t)^{c})^{2}.$$

Applying Green's identity (as in Theorem 6D) we find that

$$\tfrac{1}{2}\frac{d}{dt}\int_{M} \sum_{c} (\rho(W_t)^{c})^{2} {*1} = -\int_{M} (\operatorname{grad}(W_t)^{2} {*1},$$

and the conclusion follows:

THEOREM. *Let $W_t \colon M \to N$ be a smooth deformation satisfying $(\bar{2})$ for $t_0 \leqq t < t_1$. If W_{t_0} maps M into M', then so does every $W_t (t_0 \leqq t < t_1)$.*

8. Derivative bounds for the elliptic case.

(A) Under suitable curvature and metric restrictions on M' we shall establish derivative estimates important for the solution of (1) and (2), or equivalently, of $(\bar{1})$ and $(\bar{2})$. Our starting point is the following result, essentially established in § 3C.

LEMMA. *Any solution f_t of (2) has energy density $e(f_t)$ satisfying*

$$\begin{aligned} \Delta e(f_t) - \frac{\partial e(f_t)}{\partial t} = {} & |\beta(f_t)|^{2} - R'_{\alpha\beta\gamma\delta} f_{ti}{}^{\alpha} f_{tj}{}^{\beta} f_{tk}{}^{\gamma} f_{tl}{}^{\delta} g^{ik} g^{jl} \\ & - g'_{\alpha\beta} f_{ti}{}^{\alpha} f_{tj}{}^{\beta} R^{ij}, \end{aligned} \tag{6}$$

where $\beta(f_t)$ is the fundamental form of f_t and where $R'_{\alpha\beta\gamma\delta}$ and R^{ij} are the components of the Riemannian curvature tensor on M' and of the contravariant Ricci tensor on M, respectively.

(B) For the elliptic systems (1), $(\bar{1})$ we invoke Green's formula for the operator Δ on M. Let $r(P,Q)$ denote the geodesic distance between points P, Q of M. Since M is compact (without boundary), there is a constant $a > 0$ such that $r^{2}(P,Q)$ is of class C^{∞} for $r^{2}(P,Q) < 3a$. Let $\phi(\lambda)$ be a monotone C^{∞} function in $0 \leqq \lambda < \infty$ with $\phi(\lambda) = \lambda$ in $0 \leqq \lambda \leqq a$ and $\phi(\lambda) =$ constant for $\lambda > 2a$. Set $\rho^{2}(P,Q) = \phi(r^{2}(P,Q))$; this is positive for $P \neq Q$ and is of class C^{∞}. The function

$$F(P,Q) = \kappa \cdot \rho(P,Q)^{-n+2} \qquad \left(= \frac{-1}{2\pi} \log \rho(P,Q) \text{ for } n = 2\right)$$

is a parametrix for the operator Δ, where $1/\kappa$ is $(n-2)$ times the surface of the unit $(n-1)$-sphere. Green's formula is

$$u(P) = \int_{M} [u(Q) \cdot \Delta_{Q} F(P,Q) - F(P,Q) \cdot \Delta u(Q)] {*1_{Q}}, \tag{7}$$

this holding for any function $u(P)$ of class C^2 on M (see Giraud [12], Bidal-de Rham [1]).

Consider now a solution of (1), and suppose that the Riemannian curvature of M' is non-positive. From (6), in which it is merely necessary to suppress the term $\partial e(f_t)/\partial t$, there follows at once

$$e(f) \geqq -\mathrm{R}e(f),$$

where R is a constant (independent of the solution in question). Since $F(P,Q) \geqq 0$ (a and ϕ being suitably chosen for $n = 2$), (7) applied to $e(f)$ yields

$$e(f)(P) \leqq \int_M [\Delta_Q F(P,Q) + R \cdot F(P,Q)] e(f)(Q) {*1_Q}.$$

By using the osculating Euclidean metric at a point of M one can show that $|\Delta_Q F(P,Q)| \leqq \text{const.} \times \rho(P,Q)^{-n+2}$. Hence for some constant A,

$$e(f)(P) \leqq A \int_M F(P,Q) e(f)(Q) {*1_Q}.$$

Iterating this $k-1$ times we obtain

$$e(f)(P) \leqq A^k \int_M F_k(P,Q) e(f)(Q) {*1_Q}.$$

where the F_k are defined inductively by $F_1 = F$ and

$$F_k(P,Q) = \int_M F_{k-1}(P,Z) \cdot F(Z,Q) {*1_Z} \qquad (k > 1).$$

If $k > n/2$, then F_k is bounded (see Giraud [12]), and we have the following

THEOREM. *If M' has non-positive Riemannian curvature, then there is a constant C such that $e(f) < C \cdot E(f)$ for any harmonic mapping $f\colon M \to M'$.*

(C) Green's function $G(P,Q)$ for Δ can be written in the form $G(P,Q) = F(P,Q) + F'(P,Q)$, where F' is of class C^∞ for $P \neq Q$ and has a singularity of order lower than that of F for $P = Q$ (Giraud [12], Bidal-de Rham [1]). $G(P,Q)$ is symmetric and of class C^∞ for $P \neq Q$ and satisfies $\Delta_P G(P,Q) = \Delta_Q G(P,Q) = V^{-1}$, where V is the volume of M. Green's formula (7), with G in place of F, is

$$(8) \qquad u(P) = V^{-1} \int_M u(Q) {*1} - \int_M G(P,Q) \cdot \Delta u(Q) {*1_Q}.$$

Now let $\bar{U}$ be a compact coordinate neighborhood on M, U its interior, and let $P \in U$ have coordinates $(x^1, \cdots, x^n)$. Write

$$G_i(P,Q) = \partial G(P,Q)/\partial x^i; \quad G_{i,j}(P,Q) = \partial^2 G(P,Q)/\partial x^i \partial x^j$$

Using normal coordinates on M one can show that there is a constant C such that

$$\begin{aligned} &|G_i(P,Q)| \leqq C\cdot r(P,Q)^{-n+1} \\ (9)\qquad &|G_{i,j}(P,Q)| \leqq C\cdot r(P,Q)^{-n} \\ &|G_i(P,Z) - G_i(Q,Z)| \leqq C\cdot r(P,Q)^{\alpha}[r(P,Z)^{-n+1-\alpha} + r(Q,Z)^{-n+1-\alpha}] \end{aligned}$$

for $P, Q \in U$, α being an arbitrary but fixed number with $0 < \alpha < 1$.

If u is a solution of $\Delta u = f$ on M for some function f, then from (8) we have (for $P \in U$)

$$u_i(P) = -\int_M G_i(P,Z)\cdot f(Z)*1_Z.$$

Hence

$$|u_i(P) - u_i(Q)| \leqq \int_M |G_i(P,Z) - G_i(Q,Z)|\cdot|f(Z)|*1_Z$$

for $P, Q \in U$. Using the last inequality of (9) we obtain

$$(10)\qquad |u_i(P) - u_i(Q)|\cdot r(P,Q)^{-\alpha} \leqq C' \sup_M |f|,$$

where

$$C' = C\cdot\sup\int_M [r(P,Z)^{-n+1-\alpha} + r(Q,Z)^{-n+1-\alpha}]*1_Z.$$

Suppose now that f is Hölder-continuous with exponent α and Hölder modulus $M_\alpha(f)$. I. e., $M_\alpha(f) = \sup_M |f(P) - f(Q)|\cdot r(P,Q)^{-\alpha}$. It is a classical result of potential theory that the function

$$\int_M G_i(P,Z)\cdot[f(P_0) - f(Z)]*1_Z \qquad (P, P_0 \in U)$$

has a derivative with respect to x^j at the point P_0, given by

$$\int_M G_{i,j}(P_0,Z)\cdot[f(P_0) - f(Z)]*1_Z,$$

and by (9), this integral is majorized by

$$C\cdot M_\alpha(f)\cdot\int_M r(P_0,Z)^{-n+\alpha}*1_Z.$$

On the other hand, the function $\phi(P) = \int_M G(P,Z)*1_Z$ is a constant, since $\Delta\phi = 0$, and so $\int_M G_i(P,Z)\cdot f(P_0)*1_Z = 0$. We conclude that

$$(11) \qquad |\partial^2 u/\partial x^i \partial x^j| \leqq C' \cdot M_\alpha(f) \text{ in } U.$$

(This is an interior estimate of the type given by Hopf and Schauder (Hopf [14], Schauder [26], Miranda [18]), but differs from them in that the magnitude of u is not involved.)

(D) We apply these results to a harmonic mapping $f\colon M \to M'$. Denote the right member of $(\tilde{1})$ by F^c:

$$F^c = \pi_{ab}{}^c W_i{}^a W_j{}^b g^{ij}.$$

At this point *we impose some boundedness conditions on the embedding of M' in $\boldsymbol{R}^q$, conditions which are automatically fulfilled if M' is compact. Namely we assume that*

$$(12) \qquad \begin{gathered} |\pi_{ab}{}^c| \leqq C_0, \quad |\partial \pi_{ab}{}^c / \partial w^d| \leqq C_0 \quad \text{on } M', \\ A_1 ds_0{}^2 \leqq ds'^2 \leqq A_2 ds_0{}^2, \end{gathered}$$

where C_0, A_1, A_2 denote positive constants and where $ds_0{}^2$ denotes the line element induced on M' by the usual metric in $\boldsymbol{R}^q$.

Again let U be the interior of a compact coordinate neighborhood on M, and let P, Q be points of U. From (12) and an elementary calculation involving the Schwarz inequality for quadratic forms and the equality $e(f)P = g''_{ab} W_i{}^a W_j{}^b g^{ij}$,

$$(13) \qquad \begin{aligned} &|F^c(P) - F^c(Q)| \cdot r(P,Q)^{-\alpha} \\ &\qquad \leqq B \cdot [\bar{e} + \bar{e}^{\frac{1}{2}} \sup_{a,i} |W_i{}^a(P) - W_i{}^a(Q)| \cdot r(P,Q)^{-\alpha}], \end{aligned}$$

where B is a constant and $\bar{e} = \bar{e}(f) = \sup\{e(f)(P) : P \in M\}$. From (1) and (10),

$$|W_i{}^a(P) - W_i{}^a(Q)| \cdot r(P,Q)^{-\alpha} \leqq C' \sup_M |F^a|,$$

and plainly $|F^a| \leqq$ const. $\times\, \bar{e}$, in virtue of (12). From the compactness of M there follows the estimate

$$(14) \qquad M_\alpha(F^c) \leqq B'(\bar{e} + \bar{e}^{\frac{3}{2}})$$

for the Hölder-modulus of F^c, B' denoting a constant. Referring to (11), we have the

THEOREM. *Suppose that M' satisfies the embedding conditions* (12). *Let $U(x^1, \cdots, x^n)$ be the interior of a compact coordinate neighborhood on M. Then there is a constant C such that*

$$|\partial^2 W^a / \partial x^i \partial x^j| \leqq C(\bar{e}(f) + \bar{e}(f)^{\frac{3}{2}}) \text{ in } U, \ (1 \leqq a \leqq q),$$

for any harmonic mapping $f\colon M \rightarrow M'$, *where* $\bar{e}(f) = \sup\{e(f)(P) : P \in M\}$.

Remark. We point out that second derivative estimates can be obtained from linear theory in another way. Namely, if we write our equation $(\tilde{1})$ in the form

$$\Delta W^c + A_b{}^{cj} W_j{}^b = 0, \text{ where } A_b{}^{cj} = -\pi_{ab}{}^c W_i{}^a g^{ij},$$

then we have a linear system with bounded coefficients (in compact coordinate neighborhoods), by Theorem 8B. A Hölder-modulus for the $W_i{}^c$ will then give us a Hölder-modulus for the coefficients of the linear system, and we can apply Theorem 1 of Douglis and Nirenberg [6] to deduce second derivative estimates and analogous estimates on all higher derivatives as well. The second derivative estimates we obtain here are somewhat sharper, in certain respects; i.e., they do not involve *a priori* estimates on the magnitudes of the solutions W^c.

9. Bounds for the parabolic case.

(A) We now embark upon some analogous computations for the operator $L = \Delta - \partial/\partial t$. The function

$$K(P,Q,t) = (2\sqrt{\pi})^{-n} t^{-n/2} \exp(-\rho^2(P,Q)/4t) \tag{15}$$

is a parametrix for the operator L (ρ^2 as in the preceding paragraph). Put

$$N_1(P,Q,t) = L_P K(P,Q,t) = (\Delta_P - \partial/\partial t) K(P,Q,t)$$

and

$$N_k(P,Q,t) = \int_0^t d\tau \int_M N_{k-1}(P,Z,t-\tau) \cdot N(Z,Q,\tau) {*1_Z} \qquad (k > 1).$$

It is well known that there exists a fundamental solution H for the heat operator L on any compact Riemannian manifold M, which can be expressed in the form

$$H(P,Q,t) = K(P,Q,t) + \int_0^t d\tau \int_M K(P,Z,t-\tau) \cdot N(Z,Q,\tau) {*1_Z}, \tag{16}$$

where

$$N(P,Q,t) = \sum_{k=1}^{\infty} N_k(P,Q,t).$$

(See Milgram-Rosenbloom [17], Pogorzelski [24]). The function $H(P,Q,t)$ is symmetric in P, Q and is positive. It is of class C^∞ except for $P = Q$, $t = 0$;

and it satisfies $L_P H(P,Q,t) = L_Q H(P,Q,t) = 0$. Its spectral decomposition is

$$H(P,Q,t) = V^{-1} + \sum_{\nu=1}^{\infty} \exp(-\lambda_\nu t)\phi_\nu(P)\phi_\nu(Q),$$

where the λ_ν are the non-zero eigenvalues of Δ, the $\phi_\nu(P)$ being the corresponding orthonormal eigenfunctions.

Green's formula (analogous to 8) is

$$\begin{aligned} u(P,t) = &-\int_{t_0}^{t} d\tau \int_M H(P,Q,t-\tau)\cdot Lu(Q,\tau)*1_Q \\ &+\int_M H(P,Q,t-t_0)\cdot u(Q,t_0)*1_Q \quad (t_0 < t < t_1), \end{aligned} \tag{17}$$

this holding for any function $u(P,t)$ on M which is of class C^2 in P and C^1 in t for $t_0 \leqq t < t_1$.

(B) Suppose now that we have a solution $f_t\colon M \to M'$ of (2) defined in $0 < t < t_1$, and let M' have non-positive Riemannian curvature. According to Lemma 8A we have then $Le(f_t) \geqq -Re(f_t)$, R being a positive constant (independent of the solution in question). Since $H > 0$, there follows from (17)

$$e(f_t)(P) \leqq R\int_{t_0}^{t} d\tau \int_M H(P,Q,t-\tau)\cdot e(f_\tau)(Q)*1_Q + e_0(f_t)(P) \tag{18}$$

where

$$e_0(t_t)(P) = \int_M H(P,Q,t-t_0)e(f_{t_0})(Q)*1_Q \tag{19}$$

and $0 < t_0 < t < t_1$. Iterating (18) $k-1$ times we obtain

$$\begin{aligned} e(f_t)(P) \leqq\ & R^k \int_{t_0}^{t} d\tau \int_M H_k(P,Q,t-\tau)e(f_\tau)(Q)*1_Q \\ &+ e_0(f_t)(P) + \sum_{\nu=1}^{k-1} R^\nu \int_{t_0}^{t} d\tau \int_M H_\nu(P,Q,t-\tau)e_0(f_\tau)(Q)*1_Q, \end{aligned} \tag{20}$$

where the H_k are defined by $H_1 = H$ and

$$H_k(P,Q,t) = \int_0^t d\tau \int_M H_{k-1}(P,Z,t-\tau)H(Z,Q,\tau)*1_Z \qquad (k > 1).$$

From the integral representation it can be shown (see Pogorzelski [24]) that

$$H(P,Q,t) \leqq \text{const.} \times t^{-\alpha} r(P,Q)^{-n+2\alpha} \qquad (0 \leqq t \leqq 1).$$

where α is an arbitrary but fixed positive number less than 1. Therefore H_k is bounded for $k > n/2$ $(0 \leqq t \leqq 1, k > 1)$.

Consider now our solution f_t for $t > 1$. Putting $t - 1$ for t_0 in (20) we have

$$e(f_t)(P) \leqq \text{const.} \int_{t-1}^{t} d\tau \int_M e(f_\tau)(Q) *1 \\ + \sup_M e_0(f_\tau)(P) \cdot [1 + \sum_{\nu=1}^{k-1} R^\nu \int_0^1 d\tau \int_M H_\nu(P, Q, \tau) *1_Q]. \tag{21}$$

For the case at hand, e_0 is given from (19) by

$$e_0(f_t)(P) = \int_M H(P, Q, 1) \cdot e(f_{t-1})(Q) *1_Q.$$

Since $H(P, Q, 1)$ is bounded, we have $e_0(f_t)(P) \leqq \text{const.} \int_M e(f_{t-1})(Q) *1_Q$. Recalling that $\int_M e(f_t)(Q)$ is a decreasing function, we obtain finally from (21)

$$e(f_t)(P) \leqq \text{const.} \int_M e(f_{t-1})(Q) *1 \qquad (t > 1).$$

Any smaller value can be put in for $t - 1$ on the right, for example zero if $e(f_t)(P)$ is continuous at $t = 0$.

Making that assumption, we now obtain an estimate for the range $0 \leqq t \leqq 1$. In (20) we now put $t_0 = 0$, getting

$$e(f_t)(P) \leqq \text{const.} [\int_0^t d\tau \int_M e(f_\tau)(Q) *1 + \bar{e}_0(f_t)], \tag{22}$$

where now we have

$$e_0(f_t)(P) = \int_M H(P, Q, t) e(f_0)(Q) *1_Q,$$

and where $\bar{e}_0(f_t) = \sup\{e_0(f_t)(P) : P \in M\}$; we define $\bar{e}(f_t)$ similarly. But this function is precisely the solution of $Le_0 = 0$ that reduces at $t = 0$ to $e(f_0)(Q)$. From general principles it follows that $e_0(f_t)(P) \leqq \bar{e}(f_0)$ for $t \geqq 0$, and so (22) gives at once

$$e(f_t)(P) \leqq \text{const.}\, \bar{e}(f_0) \qquad (0 \leqq t \leqq 1).$$

We have then the

THEOREM. *Let $f_t : M \to M'$ be a family of mappings for $0 \leqq t < t_1$ satisfying* (2) *for $0 < t < t_1$ and such that the energy density $e(f_t)(P)$ is* *continuous at* $t=0$. *Suppose that* M' *has non-positive Riemannian curvature. Then*

$$e(f_t)(P) \leqq C\cdot\int_M e(f_0)(Q)*1 \qquad \textit{for } 1\leqq t<t_1$$

and

$$e(f_t)(P) \leqq C\cdot \sup\{e(f_0)(Q) : Q\in M\} \qquad \textit{for } 0\leqq t\leqq 1,$$

C *denoting a constant which does not depend on the particular solution* f_t *of* (5).

Remark. Under certain circumstances much sharper estimates can be obtained. With the hypotheses of the preceding theorem, assume further that the Ricci tensor of M is positive definite at every point. From (6) it is clear that $\partial e(f_t)/\partial t \leqq -Ae(f_t)$ at any maximum point of $e(f_t)$ on M, A denoting a positive constant. It follows easily that $e(f_t)(P)\leqq \text{const.}\, e^{-At}$.

(C) Now let $(x^1,\cdots,x^n)$ be the coordinates of P in the interior U of a compact coordinate neighborhood $\bar{U}$ on M. And suppose that our solution of (2) and the first-order space derivatives are continuous at $t=0$. Then from $(\bar{2})$, (17) we have

$$\begin{aligned} & W_t^c(P) = W^c(P,t) \\ (23) \qquad & = -\int_0^t d\tau \int_M H(P,Q,t-\tau)\cdot F^c(Q,\tau)*1_Q + W_0^c(P,t) \\ & = V^c(P,t) + W_0^c(P,t), \end{aligned}$$

where

$$W_0^c(P,t) = \int_M H(P,Q,t)W^c(Q,0)*1_Q,$$

the $F^c(P,t)$ being the functions on the right of $(\bar{2})$.

The first integral $V^c(P,t)$ has Hölder continuous first-order space derivatives (Pogorzelski [23], Theorem 5):

$$|V_i^c(P,t)-V_i^c(P',t')| \leqq \text{const.}\sup|F^c|\cdot[r(P,P')^\alpha + |t-t'|^{\alpha/2}],$$

α being an arbitrary positive number less than 1, the points P and P' both in U. If we continue with the assumption (12), we shall have $|F^c|\leqq \text{const.}\times \bar{e}$. The integral $W_0^c(P,t)$ can be differentiated under the integral sign (for $t>0$) and the derivatives tend exponentially to zero, as is quickly seen from the spectral formula for H. Hence, if the hypotheses of Theorem 9B hold, it follows that the functions $W^c(P,t)$ have first-order space derivatives which are Hölder-continuous in P, t, uniformly so for $t \geqq \epsilon > 0$:

$$\begin{aligned} &| W_i^c(P,t) - W_i^c(P',t') | \\ &\qquad \leqq \text{const.}\, [\bar{e}(f_0) + \sup_M | W^c(Q,0) |] \cdot [r(P,P')^\alpha + | t - t' |^{\alpha/2}]. \end{aligned}$$

(The constant depends upon ϵ because of the behavior of $W_0^c(P,t)$ for small t.)

Referring to (13) we see that

$$\begin{aligned} &| F^c(P,t) - F^c(Q,t) | \cdot r(P,Q)^{-\alpha} \\ &\qquad \leqq \text{const.}\, \bar{e}(f_0)\, [1 + \bar{e}(f_0)^{\frac{1}{2}} + \sup_M | W^c(Q,0) |] \end{aligned} \tag{24}$$

for $t \geqq \epsilon > 0$.

Now from (17) it is clear that $\int_0^t d\tau \int_M H(P,Q,t-\tau)\psi(\tau) * 1_Q$ is a function of t alone for arbitrary ψ. Hence for the second derivatives $V_{i,j}^c(P,t) = \partial^2 V^c(P,t)/\partial x^i \partial x^j$ we can write

$$V_{i,j}(P,t) = \frac{\partial^2}{\partial x^i \partial x^j} \int_0^t d\tau \int_M H(P,Q,t-\tau)\, [F^c(P_0,\tau) - F^c(Q,\tau)] * 1_Q.$$

From (Pogorzelski [24], Th. 3) this is

$$\begin{aligned} V_{i,j}^c(P,t) &= \int_0^t d\tau \int_M H_{i,j}(P,Q,t-\tau)\, [F^c(P_0,\tau) - F^c(Q,\tau)] * 1_Q \\ &= \int_{t-\epsilon}^t + \int_0^{t-\epsilon} = I_1 + I_2, \end{aligned}$$

where we assume $t \geqq 2\epsilon > 0$. The integral I_1 is improper but uniformly convergent. Now $| H_{i,j}(P,Q,t) | < \text{const.} \times t^{-\beta} r(P,Q)^{-n-2+2\beta}$ (arbitrary β, $0 < \beta < 1$). Using (24) and putting $P = P_0$ in I_1 (we assume $(P, P_0 \in U)$, we obtain an absolutely convergent integral if α, β are chosen properly, and there results

$$| I_1 | < \text{const.}\, \bar{e}(f_0)\, [1 + \sup_M | W^c(Q,0) | + \bar{e}(f_0)^{\frac{1}{2}}] \text{ if } t \geqq 2\epsilon,$$

since then τ in I_1 will be $\geqq \epsilon$. For I_2 we have $| H_{i,j}(P,Q,t) | < \text{const.}\, e^{-\gamma t}$ for some positive γ and for $t \geqq \epsilon > 0$. Hence

$$\begin{aligned} | I_2 | &< \text{const.} \sup | F^c | \cdot \int_0^{t-\epsilon} e^{-\gamma(t-\tau)}\, d\tau \\ &< \text{const.} \sup | F^c | < \text{const.}\, \bar{e}(f_0), \end{aligned}$$

using Theorem 9B.

THEOREM. *Suppose that M' satisfies the imbedding conditions* (12). *Let f_t satisfy the conditions of Theorem* 9B, *and let $(x^1,\cdots,x^n)$ be the coordinates of a point P in the interior of a compact coordinate neighborhood $\bar{U}$ on M. Given $\epsilon > 0$, there is a constant C, independent of the solution f_t of* (2), *such that*

$$\left|\frac{\partial^2 W^c(P,t)}{\partial x^i \partial x^j}\right| < C\cdot \bar{e}(f_0)[1+\bar{e}(f_0)^{\frac{1}{2}}+\sup\{|W^c(Q,0)|: Q\in M\}$$

for $t \geqq \epsilon$, where $\bar{e}(f_0) = \sup\{e(f_0)(P): P \in M\}$.

Remark. Since $\int_M H(P,Q,t)*1_Q$ is a constant, the functions $W^c(Q,0)$ appearing in the foregoing estimates can be altered by arbitrary additive constants without affecting the validity of the estimates. For example, one could replace $W^c(Q,0)$ by the function minus its average value, say $\hat{W}^c$, with the result that the term $\sup|W^c(Q,0)|$ in Theorem 9C would be replaced by $\sup|W^c(Q,0)-\hat{W}^c|$.

10. Successive approximations.

(A) Let $W(P,t)$ and $W'(P,t)$ be two solutions of $(\tilde{2})$ in $0 \leqq t < t_1$, both continuous along with their first order space derivatives at $t=0$; and suppose that $W(P,0) = W'(P,0) \in M'$ for all $P \in M$. From (17) with $t_0 = 0$,

$$\begin{aligned} W^c(P,t) &- W'^c(P,t) \\ &= -\int_0^t d\tau \int_M H(P,Q,t-\tau)[F^c(Q,\tau)-F'^c(Q,\tau)]*1_Q, \end{aligned}$$

where F^c, F'^c are the respective right members of $(\tilde{2})$. Set

$$\begin{aligned} X(t) = \sup_{M,c} &|W^c(P,t)-W'^c(P,t)| \\ &+ [\sup_{M,c}(W_i{}^c - W_i{}'^c)(W_j{}^c - W_j{}'^c)g^{ij}]^{\frac{1}{2}}. \end{aligned}$$

From the constitution of F^c and F'^c it is easily verified (by an argument similar to that for (13)), account taken of (12), that

$$|F^c - F'^c| < \text{const.}\, X(t)\cdot U(t),$$

where

$$U(t) = \bar{e}(f_t) + \bar{e}(f'_t) + \bar{e}(f_t)^{\frac{1}{2}} + \bar{e}(f'_t)^{\frac{1}{2}}.$$

For $0 \leqq t \leqq 1$ we can write $H(P,Q,t) < \text{const.}\, t^{-\alpha} r(P,Q)^{-n+2\alpha}$ and

$$|\partial H(P,Q,t)/\partial x^i| < \text{const.}\, t^{-\alpha} r(P,Q)^{-n-1+2\alpha} \qquad (0<\alpha<1),$$

where in the latter the constant depends upon the particular choice of local coordinates x^i, of course. Now let A denote an upper bound for the quantity $U(t)$ in some fixed time interval $0 \leqq t \leqq t_2$. From the integral expression above there follows easily, for $\frac{1}{2} < \alpha < 1$,

$$X(t) < \text{const.} \times A \int_0^t (t-\tau)^{-\alpha} X(\tau)\, d\tau,$$

and we conclude that $X(t)$ vanishes for small t. Hence the

THEOREM. *Let f_t and f'_t be two solutions of* (2), *both continuous along with their first-order space derivatives at $t=0$. If $f_0 = f'_0$, then the two solutions coincide for all* (*relevant*) $t > 0$.

COROLLARY. *Any solution of* (2) *enjoys the semi-group property along the trajectory of each point $P \in M$. That is, if we write $f_t(P) = T_t(f\ )$, then*

$$T_{t+T}(f_0) = T_t(f_T) = T_t(T_T(f_0)).$$

(B) For the solution of (2) we now turn to the system of non-linear integro-differential equations (23) associated with $(\tilde{2})$.

Let $f\colon M \to M'$ be a mapping of class C^1, given by global mapping functions $W = (W^1, \cdots, W^q)$. For $\nu \geqq 0$ define $W^\nu = (W^{\nu,1}, \cdots, W^{\nu,q})$ by

$$W^{0,c}(P,t) = \int_M H(P,Q,t)\, W^c(Q) * 1_Q \tag{25}$$

and

$$\begin{aligned} &W^{\nu,c}(P,t) \\ &\quad = -\int_0^t d\tau \int_M H_i(P,Q,t-\tau) \cdot F^{\nu-1,c}(Q,\tau) * 1_Q + W_i^{0,c}(P,t), \end{aligned} \tag{26}$$

where

$$F^{\nu,c}(P,t) = \pi_{ab}{}^c(W^\nu) \cdot W_i{}^{\nu,a} W_j{}^{\nu,b} g^{ij},$$

the functions $\pi_{ab}{}^c$ as in $(\tilde{2})$. Set

$$y_\nu = \sup_M \Big[\sum_{c=1}^{q} W_i{}^{\nu,c} W_j{}^{\nu,c} g^{ij} \Big]^{\frac{1}{2}}.$$

From (26)

$$\begin{aligned} &W_i{}^{\nu,c}(P,t) \\ &\quad = -\int_0^t d\tau \int_M H(P,Q,t-\tau) \cdot F^{\nu-1,c}(Q,\tau) * 1_Q + W^{0,c}(P,t), \end{aligned} \tag{27}$$

where the subscript i denotes differentiation with respect to a system of local coordinates x^i at P. We recall the estimate

$$|H_i(P,Q,t)| < At^{-\alpha}r(P,Q)^{-n-1+2\alpha} \qquad (0 \leqq t \leqq 1; \tfrac{1}{2} < \alpha < 1),$$

A denoting a constant which depends in general on the local coordinate system. Now let B denote an upper bound for the quantities $|\pi_{ab}{}^c|$ in some compact neighborhood U' of the image $f(M)$ in the tubular neighborhood $N \subset \mathbf{R}^q$. If $W^{\nu-1} \in U'$ for $0 \leqq t \leqq \epsilon$, then from (27) we conclude that

$$y_\nu < BC \int_0^t (t-\tau)^{-\alpha} y^2_{\nu-1}(\tau)\, d\tau + y_0(t) \qquad (0 \leqq t \leqq \epsilon),$$

where C is a constant which does not depend upon the given mapping f. Put

$$\bar{y}_\nu = \sup y_\nu, \qquad 0 \leqq t \leqq \epsilon.$$

Then

$$\bar{y}_\nu < K\epsilon^{1-\alpha}\bar{y}_{\nu-1}{}^2 + \bar{y}_0, \qquad K = BC/(1-\alpha).$$

If $K\epsilon^{1-\alpha}\bar{y}_{\nu-1} \leqq \frac{1}{2}$ and $K\epsilon^{1-\alpha}\bar{y}_0 \leqq \frac{1}{4}$, then

$$K\epsilon^{1-\alpha}\bar{y}_\nu \leqq (K\epsilon^{1-\alpha}\bar{y}_{\nu-1})^2 + K\epsilon^{1-\alpha}\bar{y}_0 \leqq \tfrac{1}{2}.$$

From a transparent induction it follows that, for sufficiently small positive ϵ, we shall have

$$W^\nu(P,t) \in U' \text{ and } \bar{y}_\nu \leqq \tfrac{1}{2}K\epsilon^{1-\alpha} \text{ for } 0 \leqq t \leqq \epsilon,\ \nu = 0, 1, 2, \cdots.$$

In particular, the defining equation (26) makes sense for all ν, provided $0 \leqq t \leqq \epsilon$.

Now put

$$X_\nu(t) = \sup_{M,c} |W^{\nu,c}(P,t) - W^{\nu-1,c}(P,t)| + [\sup_{M,c} (W_i^{\nu,c} - W_i^{\nu-1,c})(W_j^{\nu,c} - W_j^{\nu-1,c}) g^{ij}]^{\frac{1}{2}}.$$

From the definition above,

$$F^{\nu,c} - F^{\nu-1,c} = \pi_{ab}{}^c(W^{\nu-1})(W_i^{\nu,a}W_j^{\nu,b} - W_i^{\nu-1,a}W_j^{\nu-1,b})g^{ij} + [\pi_{ab}{}^c(W^\nu) - \pi_{ab}{}^c(W^{\nu-1})] W_i^{\nu,a}W_j^{\nu,b}g^{ij}.$$

We can suppose that the constant B occurring above is also an upper bound for the quantities $|\partial\pi_{ab}{}^c/\partial w^d|$ in the neighborhood U'. The preceding formula then gives us

$$| F^{\nu,c} - F^{\nu-1,c} | < C'BX_\nu(\bar{y}_\nu + \bar{y}_{\nu-1} + \bar{y}_\nu{}^2),$$

C' a constant independent of the given mapping f. Hence,

$$| F^{\nu,c} - F^{\nu-1,c} | < C'BX_\nu[1/K\epsilon^{1-\alpha} + 1/4K^2\epsilon^{2-2\alpha}] = C''X_\nu(t).$$

From (26), (27) and the estimates cited for $H(P, Q, t)$ and $H_i(P, Q, t)$ we obtain

$$X_{\nu+1}(t) < DC'' \int_0^t X_\nu(\tau) \cdot (t-\tau)^{-\alpha}\, d\tau \qquad (0 \leqq t \leqq \epsilon),$$

where D is a new constant. Thus if

$$\bar{X}_\nu(t) = \sup X_\nu(\tau) \qquad (0 \leqq \tau \leqq t),$$

then

$$\bar{X}_\nu(t) < \left(\frac{DC''}{1-\alpha} t^{1-\alpha}\right)^\nu \bar{X}_0(t) \qquad (\nu = 1, 2, \cdots);$$

and so the series

$$\sum_0^\infty \bar{X}_\nu(t)$$

converges for all sufficiently small t. This shows that our successive approximations W^ν and their first-order space derivatives converge uniformly on M (for small t). Hence, the $F^{\nu,c}$ also converge. Set $W^c = \lim W^{\nu,c}$ and $F^c = \lim F^{\nu,c}$ ($\nu \to \infty$). Thus $W^c(P, t)$ has continuous first-order space derivatives W_i^c (for sufficiently small t, of course), and $W_i^{\nu,c} \to W_i^c$ ($\nu \to \infty$), so that

$$F^c(P, t) = \pi_{ab}{}^c(W) W_i{}^a W_j{}^b g^{ij}.$$

From (26) there follows at once

$$(28) \qquad W^c(P, t) = -\int_0^t d\tau \int_M H(P, Q, t-\tau) \cdot F^c(Q, \tau) * 1_Q + W^{0,c}(P, t).$$

We conclude that the functions W^c have Hölder-continuous first order space derivatives—uniformly Hölder-continuous for any small closed t-interval not containing zero. Therefore the functions $F^c(P, t)$ are Hölder-continuous with respect to the space variable, and so the functions W^c of (28) satisfy equation ($\tilde{2}$) for all positive t in the interval in which the successive approximations converge (see Pogorzelski [24]). The W^c are moreover visibly continuous with their first order space derivatives at $t = 0$, and $W^c(P, 0) = W^c(P)$. From Theorem 7C we recall that the point $W(P, t) = (W^1, \cdots, W^q)$ must lie on the manifold M' for all t in the interval of convergence.

THEOREM. *Let M'' be a compact subset of M'. Then, for any continuously differentiable mapping $f\colon M \to M'$ such that $f(M)$ lies in M'', there is a positive constant t_1 depending only on M'' and the energy density $e(f)$ such that (2) has a solution f_t for $0 \leqq t \leqq t_1$ which is continuous at $t = 0$, along with its first order space derivatives, and which coincides with f at $t = 0$.*

It only remains for us to inquire into the length of the t-interval for which the successive approximations converge. First of all, the upper bound B for $|\pi_{ab}{}^c|$ and $|\partial\pi_{ab}{}^c/\partial w^d|$ figuring in the foregoing proof can be taken once for all to be valid in some neighborhood U'' of M''. Therefore the constant K can be fixed, and α as well, of course. The ϵ must satisfy $K\epsilon^{1-\alpha}\bar{y}_0 \leqq \frac{1}{4}$, and to see what this entails we must look briefly into $\bar{y}_0$.

Consider then a solution u of the heat equation on M, as a map $u\colon M \to \boldsymbol{R}$. We suppose that u and its first order space derivatives are continuous at $t = 0$. The argument of Theorem 9B holds in this situation. If we put $p = g^{ij}u_iu_j$, then $p(P, t) \leqq C \sup\{p(Q, 0) : Q \in M\}$, where C is a constant depending only on M. Applying this to the functions $W^{0,c}$, we conclude that there is a constant C_1 such that $\bar{y}_0 < C_1 \cdot \bar{e}(f)$. Thus ϵ depends only on M'' and the magnitude of $e(f)$, and the same is true of the quantity C'' involved in the estimates of the X_ν. The assertion of the theorem then follows at once from those estimates.

In § 2B we described the harmonic character of C^2 maps in terms of their tension fields. For C^1 maps we have the

COROLLARY. *Let $f\colon M \to M'$ be a continuously differentiable mapping for which the energy is a minimum with respect to small variations. Then f is harmonic.*

For let f_t be the corresponding solution of (2) guaranteed by the preceding theorem. The energy $E(f_t)$ is continuous at $t = 0$ and by assumption it must be non-decreasing for small t. But $E(f_t)$ is always non-increasing, so that $dW/dt = 0$ for small t. Thus $(\bar{2})$ reduces to $(\bar{1})$.

(C) THEOREM. *Suppose that M' has non-positive Riemannian curvature, and that it satisfies the embedding restrictions (12). Then for any continuously differentiable mapping $f\colon M \to M'$ there is a unique solution f_t*

of (2), *defined for all* $t \geqq 0$, *which is continuous along with its first order space derivatives at* $t = 0$ *and which coincides with* f *at* $t = 0$.

Such a solution exists for small t, by Theorem 10B, and it is unique, by Theorem 10A. Let t_1 be the largest number such that a solution of the required sort exists for $0 \leqq t < t_1$, and supposes that t_1 is finite. From Theorem 9B it is clear that the right members of (28) cannot become unbounded for $0 \leqq t < t_1$ and consequently the images $f_t(M)$ $(0 \leqq t < t_1)$ all lie in a compact subset of M'; we recall that M' is always assumed to be complete. On the other hand, Theorem 9B shows that the energy density $e(f_t)$ remains bounded, and therefore by Theorems 10A and 10B there is a fixed positive number ϵ_1 such that any f_t can be continued as a solution of (2) into the interval $(t, t + \epsilon_1)$. This contradicts the definition of t_1.

(D) If M' is not compact, then solutions of (2) may very well become unbounded as $t \to \infty$, as in the

Example. Let M' be the manifold obtained by revolving the graph of a positive strictly decreasing smooth function $v = v(u)$ around the u-axis; let ϕ denote the revolution angle. For a map $f\colon S^1 \to M'$ of a circle S^1 parametrized by the central angle θ our heat equation is

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial \theta^2} + \frac{v'v''}{1 + (v')^2}\left(\frac{\partial u}{\partial \theta}\right)^2 - \frac{vv'}{1 + (v')^2}\left(\frac{\partial \phi}{\partial \theta}\right)^2$$

$$\frac{\partial \phi}{\partial t} = \frac{\partial^2 \phi}{\partial \theta^2} + 2\,\frac{v'}{v}\,\frac{\partial u}{\partial \theta}\,\frac{\partial \phi}{\partial \theta}.$$

If f satisfies initial conditions $\partial u/\partial\theta = 0$, $\phi = \theta$ when $t = 0$, then so does the solution f_t for any subsequent time. If we take $v(u) = 1 + e^{-u}$, then $R'_{1212} = -(e^u + 1)/(e^{2u} + 1) < 0$, and the heat equation reduces to

$$\frac{\partial u}{\partial t} = \frac{e^u + 1}{e^{2u} + 1}.$$

Thus $e^u + u - 2\log(e^u + 1) = t + \text{const.}$; in particular, $u \to \infty$ as $t \to \infty$.

We note in passing that there are no non-trivial closed geodesics on M', so that there are no harmonic representatives in any non-trivial homotopy classes of maps $S^1 \to M'$.

The following result shows that solutions must remain bounded if M' satisfies certain conditions at infinity.

THEOREM. *Let M' be as in the preceding theorem and suppose further that* $|w| \cdot \pi_{ab}{}^{c}(w) \to 0$ *uniformly as* $|w| \to \infty$, *where* $|w| = \sup\limits_{c} |w^c|$. *Then every solution of* $(\bar{2})$ *is bounded.*

Set

$$\bar{W}(t) = \sup_{M,c} |W^c(P,t)|, \quad \underline{W}(t) = \inf_{M,c} |W^c(P,t)|, \quad U(t) = \int_M \sum_c (W^c)^2 {*1}.$$

In virtue of Theorem 9B, the difference $\bar{W}(t) - \underline{W}(t)$ is bounded. Hence if our solution is unbounded as $t \to \infty$, then that is true of all three of the quantities above. Supposing that to be the case, let us denote by λ_k the set of all t for which $U(t) > k$. The λ_k are then all non-empty and each λ_k must contain at least one $t = t_k$ at which $dU/dt \geqq 0$. Now from $(\bar{2})$ we have

$$\sum_c W^c \Delta W^c - \sum_c W^c \frac{dW^c}{dt} = \sum W^c \pi_{ab}{}^{c} W_i{}^{a} W_j{}^{b} g^{ij}.$$

Hence by Green's Theorem

$$\tfrac{1}{2} dU/dt = -\int_M (\operatorname{grad} W)^2 {*1} - \sum_c \int_M W^c \pi_{ab}{}^{c} W_i{}^{a} W_j{}^{b} g^{ij} {*1}.$$

For large values of t_k we have a plain contradiction, since the right-hand side must be negative.

11. Harmonic mappings.

(A) We can now apply some of the results established above to prove the existence of harmonic mappings, even though we do not know whether the solutions of the parabolic system (2) converge in general as $t \to \infty$.

THEOREM. *Let M' have non-positive Riemannian curvature and let* $f_t \colon M \to M'$ *be a bounded solution of* (2), $0 < t < \infty$. *Then there is a sequence* $t_1, t_2, t_3, \cdots$ *of t-values such that the mappings* f_{t_k} *converge uniformly, along with their first order space derivatives, to a harmonic mapping f.*

From Theorems 9B and 9C it is clear that the mappings f_t and their first order space derivatives form equicontinuous families. Hence there exists a sequence $t_1, t_2, \cdots$ such that the mappings $f_k = f_{t_k}$ converge uniformly, with their first order space derivatives, to a continuously differentiable mapping f. From $(\bar{2})$ and (8) we can represent the f_k by the formula

$$W^c(P,t_k) = V^{-1} \int_M W^c(Q,t_k) {*1} - \int_M G(P,Q) \left[F^c(Q,t_k) + \frac{dW^c}{dt}(Q,t_k)\right] {*1_Q},$$

where as usual F^c stands for the right member of $(\bar{2})$. Now fix c and temporarily put

$$u_k(P) = \int_M G(P,Q)\,\frac{dW^c}{dt}\,(Q,t_k)*1_Q.$$

The $dW^c(P,t_k)/dt$ are bounded as $k\rightarrow\infty$, by Theorems 9B and 9C, and so the u_k and their first derivatives are bounded. Hence the u_k form an equicontinuous family, and we can suppose that the sequence $t_1, t_2, \cdots$ is chosen so that the u_k converge uniformly, say to u. Now let G_ν denote the ν-th iterate of the Green's function G. We have

$$\begin{aligned}\int_M G_\nu(P,Q)u(Q)*1 &= \lim_k \int_M G_\nu(P,Q)u_k(Q)*1\\ &= \lim_k \int_M G_{\nu+1}(P,Q)\,\frac{dW^c(Q,t_k)}{dt}\,*1.\end{aligned}$$

If $\nu+1>n/2$, then $G_{\nu+1}$ is bounded. But the $dW^c(P,t_k)/dt$ converge in the mean to zero as $k\rightarrow\infty$, by Corollary 6C. Thus

$$\int_M G_\nu(P,Q)u(Q)*1=0,$$

and so $u=0$ because of the positive-definite character of G. Therefore, passing to the limit in the equation above, we get for the limit mapping f the formula

$$W^c(P) = V^{-1}\int_M W^c(Q*1 - \int_M G(P,Q)\cdot F^c(Q)*1_Q,$$

where

$$F^c(Q) = \lim_k F^c(Q,t_k) = \pi_{ab}{}^c(W)\,W_i{}^a W_j{}^b g^{ij}.$$

From this it follows (as in § 10B) that $W^c(P)$ has Hölder-continuous first derivatives, and therefore f^c is Hölder-continuous. Consequently the $W^c(P)$ satisfy $(\bar{1})$.

COROLLARY. *Let M' have non-positive Riemannian curvature and let $f\colon M\rightarrow M'$ be a continuously differentiable mapping. Let f_t be the solution of* (2) *which reduces to f at $t=0$. If f_t is bounded as $t\rightarrow\infty$, then f is homotopic to a harmonic mapping f' for which $E(f')\leqq E(f)$. In particular, if M' is compact or satisfies the conditions of Theorem* 10D, *then every continuous mapping $M\rightarrow M'$ is homotopic to a harmonic mapping.*

COROLLARY. *If M' is compact and has non-positive Riemannian curvature, then every homotopy class of mappings $M \to M'$ contains a harmonic mapping whose energy is an absolute minimum.*

For in any homotopy class we can choose a minimizing sequence of *harmonic* mappings f_1, f_2, etc., by the preceding corollary. From Theorems 8B and 8D it follows that we can select a subsequence (same notation) which converges uniformly along with first derivatives to a continuously differentiable mapping f. Then $E(f) = \lim E(f_k)$, and f is harmonic by Corollary 10B.

(B) THEOREM. *Let M have non-negative Ricci curvature and let M' have non-positive Riemannian curvature. Suppose that M' is compact or that it satisfies the conditions of Theorem* 10D. *Then any continuous map $f\colon M \to M'$ is homotopic to a totally geodesic map. Furthermore,*

1) *if there is at least one point of M at which its Ricci curvature is positive, then every continuous map from M to M' is null-homotopic;*

2) *if the Riemannian curvature of M' is everywhere negative, then every continuous map from M to M' is either null-homotopic or is homotopic to a map of M onto a closed geodesic of M'.*

This is a combination of Theorem 3C and Corollary 11A.

COROLLARY. *Let M be a compact smooth manifold admitting a Riemann structure g' with non-positive Riemannian curvature. Then M does not admit any Riemannian structure g with non-negative Ricci curvature unless that curvature vanishes everywhere.*

Proof. It suffices (by passing to the two leaved orientable cover if necessary) to consider the case that M is orientable. If there were such a metric g, then the identity map $(M, g) \to (M, g')$ would be homotopic to a totally geodesic map. That map has degree one, and therefore M cannot have any point of positive Ricci curvature relative to g.

Remark. A special case of Theorem 11B (Part 1) can be obtained without the use of harmonic theory. Namely, assume

1) that M has positive Ricci curvature everywhere; then by a theorem of S. Myers [20], the fundamental group $\pi_1(M, m)$ is finite;

2) that M' is any complete Riemannian manifold with non-positive Riemannian curvature. Then the homotopy groups $\pi_i(M', m') = 0$ for $i \neq 1$, and $\pi_1(M', m')$ has no elements of finite order.

It is well known that the homotopy classes of maps of any arcwise connected space M into M' are in natural 1-1 correspondence with the conjugacy classes of homomorphisms $\pi_1(M, m) \to \pi_1(M', m')$. But clearly in the situation at hand every such homomorphism is trivial, whence every continuous map $M \to M'$ is null homotopic.

COLUMBIA UNIVERSITY,
THE JOHNS HOPKINS UNIVERSITY,
THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] P. Bidal and G. de Rham, "Les formes différentielles harmonique," *Commentarii Mathematici Helvetici*, vol. 19 (1946-7), pp. 1-49.

[2] S. Bochner, "Harmonic surfaces in Riemannian metric," *Transactions of the American Mathematical Society*, vol. 47 (1940), pp. 146-154.

[3] ———, "Analytic mappings of compact Riemann spaces onto Euclidean space," *Duke Mathematical Journal*, vol. 3 (1937), 339-354.

[4] ———, "Curvature and Betti numbers in real and complex vector bundles," *Rendiconti del Seminario Matematico di Torino*, vol. 15 (1955-6), pp. 225-253.

[5] C. Chevalley, *Lie Groups I*. Princeton, 1946.

[6] A. Douglis and L. Nirenberg, "Interior estimates for elliptic systems of partial differential equations," *Communications on Pure and Applied Mathematics*, vol. 8 (1955), pp. 503-538.

[7] F. Dressel, "The fundamental solution of the parabolic equation I," *Duke Mathematical Journal*, vol. 7 (1940), pp. 186-203.

[8] L. P. Eisenhart, *Riemannian Geometry*. Princeton, 1926.

[9] A. Friedman, "Interior estimates for parabolic systems of partial differential equations," *Journal of Mathematics and Mechanics*, vol. 7 (1958), pp. 393-418.

[10] F. B. Fuller, "Harmonic mappings," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 40 (1954), pp. 987-991.

[11] M. Gevrey, "Sur les équations aux dérivées partielles du type parabolique," *Journal de Mathématiques et Appliquées*, s. 6, vol. 9 (1913), pp. 305-475.

[12] G. Giraud, "Sur le problème de Dirichlet généralisé," *Annales de l'École Normale Supérieure*, vol. 46 (1929), pp. 131-245.

[13] R. Hermann, "A sufficient condition that a map of Riemannian manifolds be a fibre bundle," *Proceedings of the American Mathematical Society*, vol. 11 (1960), pp. 236-242.

[14] E. Hopf, "Über den funktionalen, insbesondere den analytischen Charakter der Lösungen elliptischer Differentialgleichungen zweiter Ordnung," *Mathematische Zeitschrift*, vol. 34 (1931), pp. 194-233.

[15] A. Lichnérowicz, "Quelques théorèmes de géométrie différentielle globale," *Commentarii Mathematici Helvetici*, vol. 22 (1949), pp. 271-301.

[16] C. Maxwell, *A Treatise on Electricity and Magnetism*, Oxford, 1892.

[17] A. Milgram and P. Rosenbloom, "Harmonic forms and heat conduction I," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 37 (1951), pp. 180-184.

[18] C. Miranda, *Equazioni alle Derivate Parziali di Tipo Ellittico*, Berlin, 1955.

[19] C. B. Morrey, "The problem of Plateau on a Riemannian manifold," *Annals of Mathematics*, vol. 49 (1948), pp. 807-851.

[20] S. Myers, "Riemannian manifolds with positive mean curvature," *Duke Mathematical Journal*, vol. 8 (1941), pp. 401-404.

[21] J. Nash, "The imbedding problem for Riemannian manifolds," *Annals of Mathematics*, vol. 63 (1956), pp. 20-64.

[22] I. Petrovsky, "On the analytic nature of solutions of systems of differential equations," *Matematičeskiy Sbornik*, vol. 47 (1939), pp. 3-70.

[23] W. Pogorzelski, "Propriétés des intégrales de l'équation parabolique normale," *Annales Polonici Mathematici*, vol. 4 (1957), pp. 61-92.

[24] ———, "Etude de la solution fondamentale de l'équation parabolique," *Ricerche di Matematica*, vol. 5 (1956), pp. 25-56.

[25] S. Sasaki, "On the differential geometry of tangent bundles of Riemannian manifolds," *Tôhoku Mathematical Journal*, vol. 10 (1958), pp. 338-354.

[26] J. Schauder, "Numerische Abschätzungen in elliptischen linearen Differentialgleichungen," *Studia Mathematica*, vol. 5 (1934), pp. 34-42.

[27] D. Spencer, Mimeographed notes, Princeton University.

ANALYTIC DISKS IN MAXIMAL IDEAL SPACES.*

By JOHN WERMER.[1]

Let X be a compact Hausdorff space. Denote by $C(X)$, resp. $C_R(X)$, the algebra of all continuous complex-valued, resp. real-valued, functions on X. A *function algebra* on X is then a closed subalgebra of $C(X)$ which contains the constants and separates the points of X. We put $\| f \| = \max_X | f |$, if f is in $C(X)$.

Let A be a function algebra on X. Denote by M the maximal ideal space of A, taken in the Gelfand topology and hence a compact Hausdorff space. We may regard X as a closed subset of M and the functions in A as defined and continuous on all of M.

Definition. Fix m in M. We say *m is the center of an analytic disk* if there exists a one-one continuous map ϕ of some open disk $|z| < r$, $r > 0$, into M such that $\phi(0) = m$ and such that for each f in A, $f(\phi)$ is analytic in $|z| < r$.

Definition. For m, m' in M, put $\| m - m' \| = \sup | f(m) - f(m') |$, taken over all f in A with $\| f \| \leq 1$. The relation $\| m - m' \| < 2$ is an equivalence relation on M (Gleason, [2]). An equivalence class is called a *part of M*.

Definition. $\operatorname{Re} A = \{\operatorname{Re} f \mid f \text{ in } A\}$. $A^{-1} = \{f \text{ in } A \mid f^{-1} \text{ is in } A\}$.

In [1] we showed the following: Assume

(1) $\operatorname{Re} A$ is dense (in the uniform norm) in $C_R(X)$.

Then if m is in M and is not a one-point part (i.e. there is some m' in M, $m' \neq m$, $\| m' - m \| < 2$), m is the center of an analytic disk.

In [3] Hoffman deduced the same conclusion from the weaker hypothesis:

(2) $\{\log | f | \mid f \text{ in } A^{-1}\}$ is uniformly dense in $C_R(X)$.

Now, roughly speaking, (1) or (2) are axioms for analytic function theory in a simply connected domain. More precisely, let Ω be a region on a Riemann surface whose boundary $\partial\Omega$ consists of finitely many analytic

---

Received May 13, 1963.

* This research was partially supported by NSF Grant GP-187.

[1] Fellow of the Alfred P. Sloan Foundation.

simple closed curves, with $\Omega \cup \partial\Omega$ compact. Let $A(\Omega)$ be the function algebra on $\partial\Omega$ consisting of all functions in $C(\partial\Omega)$ which extend continuously to $\Omega \cup \partial\Omega$ to be analytic on Ω. Then $A(\Omega)$ satisfies (1) or (2) if and only if Ω is simply connected. On the other hand, for any Ω we have:

LEMMA 1. *There exists a finite set of functions in* $A(\Omega)^{-1}$, $Z_1, \cdots, Z_n$, *such that the (real) vector space spanned by* $\operatorname{Re} A$ *and the functions* $\log|Z_i|$, $i=1, \cdots, n$ *is dense in* $C_R(\partial\Omega)$.

Proof. Let $\gamma_1, \cdots, \gamma_n$ be a homology basis of closed paths in Ω. Define functionals L_i on $C_R(\partial\Omega)$ as follows: if u is in $C_R(\partial\Omega)$, let U be the harmonic function in Ω having u as boundary values, and V the (in general multiple-valued) harmonic conjugate of U. Put

$$L_i(u) = \frac{1}{2\pi}\int_{\gamma_i} dV$$

Choose u_i in $C_R(X)$ such that u_i is twice continuously differentiable on $\partial\Omega$ and $L_i(u_i) = 1$, $L_i(u_j) = 0$, $i \neq j$. Let U_i, V_i be the corresponding harmonic functions and set $Z_j = \exp(U_j + iV_j)$. Then Z_j will have 0 periods over every γ_i, hence be single-valued on Ω. Also V_j is locally continuous on $\partial\Omega$ and so Z_j belongs to $A(\Omega)^{-1}$. On $\partial\Omega$, $\log|Z_i| = u_i$. If u is any smooth function in $C_R(\partial\Omega)$, put $\alpha_i = L_i(u)$ and consider $u' = u - \sum_{i=1}^{n} \alpha_i \log|Z_i|$. Then $L_j(u') = 0$ for all j and so u' is in $\operatorname{Re} A(\Omega)$. Since smooth functions are dense in $C_R(\partial\Omega)$, we get the assertion.

We are thus led to make the following requirement on a function algebra A on a space X:

(3) There exist Z_i in A^{-1}, $1 \leq i \leq s$, such that the subspace of $C_R(X)$ spanned by $\operatorname{Re} A$ and $\log|Z_i|$, $i = 1, \cdots, s$, is dense in $C_R(X)$.

Definition. Fix a in M. An *Arens-Singer* measure for a is a probability measure μ_a on X satisfying

$$\log|f(a)| = \int_X \log|f|\, d\mu_a, \text{ all } f \text{ in } A^{-1}.[8]$$

The existence of such a measure for every a in M was shown in [4]. Under (3), each μ_a is evidently unique.

We now fix some b in M. Assume:

(4) If dP is any real measure on X with $\int u \cdot dP = 0$ for all u in $\operatorname{Re} A$, then $dP = G \cdot d\mu_b$ for some G in $L^2(\mu_b)$.

---

[8] Since, for $f \in A$, $\operatorname{Re} f = \log|e^f|$ and $e^f \in A^{-1}$, it follows that also $\int g\, d\mu_a = g(a)$, all $g \in A$.

THEOREM 1. *Assume A satisfies* (3) *and* (4).† *Let a be any point of M lying in the part of b, with $a \neq b$. Then a is the center of an analytic disk.*

Note. Without loss of generality we may assume, for each i, $1 \leqq i \leqq s$, that $\log |Z_i|$ does not lie in the closed subspace of $C_R(X)$ spanned by $\operatorname{Re} A$ and the $\log |Z_j|$ with $j \neq i$.

Then for each i there is a real measure dP_i on X with $\int \log |Z_i| \cdot dP_i = 1$, $\int \log |Z_j|\, dP_i = 0$, $j \neq i$, and $\int u\, dP_i = 0$, u in $\operatorname{Re} A$. Because of (4) there is some G_i in $L^2(\mu_b)$ with $dP_i = G_i \cdot d\mu_b$.

Definition. If m is any point of M and μ_m its Arens-Singer measure and if F is any element of $L^1(\mu_m)$, $F(m) = \int F \cdot d\mu_m$.

Note. Let x and y lie in the same part of M. In [2] Gleason obtained several consequences for x and y. In the following remarks and in the proof of the following Lemma we make arguments closely related to Gleason's.

Put for $0 \leqq r < 1$, $K(r) = \sup |F(y)|$ taken over all F in A with $|F(x)| \leqq r$, $\| F \| \leqq 1$. Suppose $K(r) = 1$. Then $\exists\, F_n$ in A with $\| F_n \| \leqq 1$, $|F_n(x)| \leqq r$, $|F_n(y)| \rightarrow 1$. Choose conformal automorphisms ϕ_n of the unit disk such that $|\phi_n(F_n(x)) - \phi_n(F_n(y))| \rightarrow 2$. Since $\phi_n(F_n)$ is in A and $\| \phi_n(F_n) \| \leqq 1$, this contradicts $\| x - y \| < 2$. Hence $K(r) < 1$, all $r < 1$. Clearly also K increases with r.

LEMMA 2. *Let x lie in the same part as b. Then μ_x is absolutely continuous with respect to μ_b.*

Proof. Suppose μ_x is not absolutely continuous with respect to μ_b. Then there exists a set Q with $\mu_b(Q) = 0$, $\mu_x(Q) > 0$. Hence for some closed G_δ set $E \subseteq Q$, $\mu_b(E) = 0$, $\mu_x(E) > 0$.

Let U be the characteristic function of E. Choose a sequence q_n in $C_R(X)$, non-increasing and tending pointwise to U. By (3), we can find u_n in $\operatorname{Re} A$ and scalars $\lambda^i{}_n$ such that if $U_n = u_n + \sum_i \lambda^i{}_n \log |Z_i|$,

$$q_n \leqq U_n \leqq q_n + 1/n.$$

Thus $U_n \rightarrow U$ pointwise and boundedly.

Now $\lambda^i{}_n = \int U_n \cdot dP_i \rightarrow \int U\, dP_i = 0$, since $\mu_b(E) = 0$. Hence $u_n \rightarrow U$ boundedly pointwise. We may assume $u_n \geqq 0$. We have $u_n = \operatorname{Re} f_n$, f_n in A. Put $F_n = \exp(-f_n)$. Then F_n is in A, $\| F_n \| \leqq 1$,

$$|F_n(x)| = \exp(-u_n(x)) \rightarrow \exp(-U(x)) = \exp(-\mu_x(E)) = r_0 < 1.$$

---

† *Added in proof*: Mr. B. O'Neill of Brown University has shown that hypothesis (4) may be dropped in Theorem 1.

Hence $|F_n(x)| \leqq r_1 < 1$, $n > n_0$. Hence $|F_n(b)| \leqq K(r_1) < 1$, $n > n_0$. Hence $\exp(-\mu_b(E)) < 1$. This is a contradiction. So we are done.

LEMMA 3. *Let a lie in the part of b. Then $d\mu_a = K \cdot d\mu_b$, K in $L^2(\mu_b)$.*

Proof. By Lemma 2 and the Radon-Nikodym theorem, we can find K in $L^1(\mu_b)$, with $d\mu_a = K d\mu_b$, $K \geqq 0$. Suppose K is not in L^2. Then there exists a sequence u_n of functions in $C_R(X)$ such that (i) $u_n \geqq 0$, (ii) $\int u_n{}^2 \cdot d\mu_b \to 0$, and (iii) $\int K u_n \cdot d\mu_b \to 1$. Because of (3), it is no loss of generality to assume $u_n = \operatorname{Re} f_n + \sum_i \lambda_1{}^{(n)} \log |Z_i|$, with $f_n \in A$, $\lambda_1{}^{(n)}$ scalars. Then $\lambda_1{}^{(n)} = \int u_n \cdot G_i d\mu_b \to 0$, each i, and $\int u_n \cdot d\mu_b \to 0$ and $\int u_n \cdot d\mu_a = \int u_n K \, d\mu_b \to 1$. Put $F_n = \exp(-f_n)$. Then F_n is in A and

$$\| F_n \| = \max \exp(-\operatorname{Re} f_n) \leqq \| \exp(-u_n) \| (1 + o(1)) = 1 + o(\mathbf{1}).$$

$$| F_n(a) | = \exp(-u_n(a) + \sum_i \lambda_i{}^{(n)} \log |Z_i(a)|) \to e^{-1}$$

and

$$| F_n(b) | = \exp(-u_n(b) + \sum_i \lambda_i{}^{(n)} \log |Z_i(b)|) \to 1.$$

This contradicts that a and b are in the same part. Thus K is in L^2.

Definition. H^2 is the closure of A in $L^2(\mu_b)$.

H^∞ is the intersection of H^2 and $L^\infty(\mu_b)$.

LEMMA 4. *H^∞ is closed under multiplication, and is Banach algebra in the L^∞-norm. If a lies in the part of b, the map $f \to f(a)$ is a homomorphism of H^∞ into the complex field.*

Proof. Given F, G in H^∞. Fix $\epsilon > 0$. Choose f in A with $\| F - f \|_{L^2} < \dfrac{\epsilon}{\| G \|_\infty}$, then choose g in A with $\| G - g \|_{L^2} < \dfrac{\epsilon}{\| f \|_\infty}$.

Then

$$\begin{aligned} \| FG - fg \|_{L^2} &= \| (F - f)G + (G - g)f \| \\ &\leqq [\int | F - f |^2 d\mu_b]^{\frac{1}{2}} \| G \|_\infty + [\int | G - g |^2 d\mu_b]^{\frac{1}{2}} \| f \|_\infty < 2\epsilon. \end{aligned}$$

Hence $F \cdot G$ lies in the L^2-closure of A, and so is in H^∞. Also H^∞ is clearly closed in the norm of L^∞, and so is a Banach algebra in that norm. The last assertion easily follows from the preceding.

Definition. $\Xi = \{ | \Phi | \cdot | Z_1 |^{\lambda_1} \cdots | Z_n |^{\lambda_s} \mid \Phi \in H^2, \lambda_1, \cdots, \lambda_s \text{ real} \}$. The elements of Ξ are thus non-negative functions defined a. e. $- d\mu_b$ on X, and also defined at each point in M which lies in the same part as b. From now on a denotes a point satisfying the hypothesis of Theorem 1.

Definition. $\Xi_a = \{F \in \Xi \mid F(a) = 0, \int F^2 \, d\mu_b \leqq 1\}$.

LEMMA 5. *There exists W in Ξ_a maximizing $F(b)$.*

Proof. Put $\gamma = \sup F(b)$, taken over all F in Ξ_a. When λ is the s-tuple of numbers $(\lambda_1, \lambda_2, \cdots, \lambda_s)$, we shall write $|Z|^\lambda$ for $|Z_1|^{\lambda_1} \cdots |Z_s|^{\lambda_s}$.

Fix $F = |\Phi| \cdot |Z|^\lambda$ in Ξ_a. Write $\lambda_i = [\lambda_i] + \theta_i$, where $[\lambda_i]$ is the integer part of λ_i and $0 \leqq \theta_i < 1$. Then

$$F = |\Phi_1| \cdot |Z|^\theta, \quad \Phi_1 = \Phi \cdot Z_1^{[\lambda_1]} \cdots Z^{[\lambda_s]} \in H^2, \quad \theta = (\theta_1, \cdots, \theta_s).$$

Without loss of generality, we may thus write each F in Ξ_a in the form $F = |\Phi| \cdot |Z|^\lambda$, Φ in H^2 and each λ_i in $[0, 1)$.

Since each Z_i is in A^{-1}, there exists $\delta > 0$ with $|Z|^\lambda \geqq \delta$ on X for all λ with $0 \leqq \lambda_i < 1$. Hence there is a constant K with $\int |\Phi|^2 \, d\mu_b < K$ if $F = |\Phi| \cdot |Z|^\lambda$ is in Ξ_a. Thus $|\Phi(b)| = |\int \Phi \, d\mu_b| \leqq \sqrt{K}$. Also $|Z|^\lambda(b) \leqq K'$. Hence $F(b) \leqq K''$, all F in Ξ_a, whence $\gamma < \infty$.

Choose F_n in Ξ_a, $F_n(b) \to \gamma$. We may assume $F_n = |\Phi_n| \, |Z|^{\lambda^{(n)}}$ where $\lambda^{(n)} \to \lambda^{(0)}$ in R^s. Also, since $\int |\Phi_n|^2 \, d\mu_b < K$ we may assume $\Phi_n \to \Phi_0$ weakly in $L^2(\mu_b)$. Since H^2 is a closed subspace of $L^2(\mu_b)$, hence weakly closed, Φ_0 is in H^2. Define $W = |\Phi_0| \, |Z|^{\lambda_0}$. Then W is in Ξ. Also

$$\Phi_n(a) = \int \Phi_n \frac{d\mu_a}{d\mu_b} \cdot d\mu_b \to \int \Phi_0 \frac{d\mu_a}{d\mu_b} d\mu_b = \Phi_0(a).$$

Since each F_n is in Ξ_a, $\Phi_n(a) = 0$, each n, and so $\Phi_0(a) = 0$, whence $W(a) = 0$. Also $\int |\Phi_n|^2 |Z|^{2\lambda^{(n)}} d\mu_b \leqq 1$, all n. Hence $\int |\Phi_n|^2 |Z|^{2\lambda^0} d\mu_b \leqq 1 + o(1)$. It follows that $\int |\Phi_0|^2 |Z|^{2\lambda^0} d\mu_b \leqq 1$. Thus W is in Ξ_a. Finally, $W(b) = |\Phi_0(b)| \, |Z(b)|^{\lambda^0}$. Hence $W(b) = \lim F_n(b) = \gamma$.

LEMMA 6. $W = 1$ a. e. — $d\mu_b$ *on* X.

Proof. Choose U continuous on X with $\int U \, d\mu_b = 0$. By (3), we can find u_n in $\operatorname{Re} A$ and $\lambda^{(n)}$ in R^s such that $U_n = u_n + \log(|Z|^{\lambda^{(n)}})$ tends to U uniformly, and we may assume $\int U_n \cdot d\mu_b = 0$, all n.

With $u_n = \operatorname{Re} f_n$, f_n in A, put $F_n = |\exp(f_n)| \cdot |Z|^{\lambda^{(n)}}$. Then F_n is in Ξ. Also $F_n = \exp U_n$, and $F_n(b) = \exp(\int U_n \, d\mu_b) = 1$. Put $G_n = W \cdot F_n$. Then G_n is in Ξ. Clearly $G_n(a) = 0$ and $G_n(b) = W(b)$.

By the maximum property of W in Ξ_a, it follows that $\int G_n^2 \cdot d\mu_b \geqq 1$ and so $\int W^2 \cdot \exp(2U_n) \, d\mu_b \geqq 1$. Let $n \to \infty$. We get $\int W^2 \cdot \exp(2U) \, d\mu_b \geqq 1$. Replace U by tU, t a real constant. Thus $A(t) = \int W^2 \cdot \exp(2tU) \, d\mu_b \geqq 1$

$= A(0)$, all real t. Hence $(d/dt)(A) = 0$ at $t = 0$, and so $\int W^2 \cdot 2U \, d\mu_b = 0$. Thus the measure $W^2 \, d\mu_b$ annihilates U whenever $d\mu_b$ does. Hence W^2 is constant a. e., and so 1 a. e.

COROLLARY. *There exist finite positive d, D with $d < |\Phi_0| < D$* a. e. — $d\mu_b$.

Definition. With λ^0 as in Lemma 5, set $d\mu' = |Z|^{2\lambda^0} \cdot d\mu_b$.

LEMMA 7. *There exist N functions $R_1, \cdots, R_N$ in $L^2(\mu_b)$, with $N \leqq s$, such that if f is in H^∞ and $f(a) = 0$, then f is divisible by Φ_0 in the ring H^∞ if and only if $\int f \cdot \bar{R}_i \cdot d\mu' = 0$, $i = 1, \cdots, N$.*

Proof. Note that $L^2(\mu')$ and $L^2(\mu_b)$ are identical as function spaces and have equivalent norms. Put

$$\mathfrak{B} = \{F \cdot \Phi_0 \mid F \text{ is in } H^2\}.$$

If F is in H^2, there exist f_n in A, $f_n \to F$ in L^2. Then

$$\int |F\Phi_0 - f_n\Phi_0|^2 \, d\mu_b = \int |\Phi_0|^2 \, |F - f_n|^2 \, d\mu_b \to 0$$

whence $F \cdot \Phi_0$ is in H^2. Thus $\mathfrak{B}$ is contained in H^2. Also, since $|\Phi_0| \geqq d$, $\mathfrak{B}$ is a closed subspace of H^2, and also a closed subspace of $L^2(\mu')$. Let $H_a^2 = \{f \text{ in } H^2 \mid f(a) = 0\}$. Then $\mathfrak{B}$ is contained in H_a^2.

Write $H_a \theta \mathfrak{B}$ for the orthocomplement of $\mathfrak{B}$ in H_a^2 in the inner product of $L^2(\mu')$. Fix k in $H_a^2 \theta \mathfrak{B}$. Since $f\Phi_0$ is in $\mathfrak{B}$, all f in A, we have

$$\int f \Phi_0 \bar{k} \, d\mu' = 0, \text{ all } f \text{ in } A. \tag{5}$$

Fix f in A, $f(b) = 0$. Then fk is in H^2 and $fk(a) = 0$, $fk(b) = 0$. Put $W_t = |\Phi_0 + t \cdot f \cdot k| \cdot |Z|^{\lambda^0}$. Then W_t is in Ξ for every complex t. Also $W_t(b) = W(b)$, $W_t(a) = 0$. Hence

$$\int W_t^2 \, d\mu_b \geqq 1, \text{ or } \int |\Phi_0 + t \cdot f \cdot k \cdot|^2 \, d\mu' \geqq \int |\Phi_0|^2 \, d\mu',$$

all scalars t. Hence $\Phi_0 \perp fk$ in $L^2(\mu')$, i. e.

$$\int \Phi_0 \overline{fk \, d\mu'} = \int \bar{f} (\Phi_0 \bar{k}) \, d\mu' = 0, \ f \in A, \ f(b) = 0. \tag{6}$$

Because of (5) and (6), $\Phi_0 \bar{k} \, d\mu'$ annihilates $\operatorname{Re} A$.

Consider now the orthocomplement $(\operatorname{Re} A)^\perp$ of $\operatorname{Re} A$ in $L^2(\mu_b)$. $(\operatorname{Re} A)^\perp$ has dimension s. Since $\Phi_0 \bar{k} \, d\mu'$ annihilates $\operatorname{Re} A$, $\bar{\Phi}_0 \cdot k \cdot |Z|^{2\lambda_0}$ is in $(\operatorname{Re} A)^\perp$. This holds for each k in $H_a^2 \theta \mathfrak{B}$. Hence $H_a^2 \theta \mathfrak{B}$ has dimension $N \leqq s$.

Let now f be an element of H^∞ with $f(a) = 0$. Then f is divisible by Φ_0 in H^∞ *if* and only if f is in $\mathfrak{B}$, if and only if f is orthogonal to $H_a^2 \theta \mathfrak{B}$, and so at last if and only if $\int f\cdot\bar{R}_i\,d\mu'=0$, $i=1,\cdots,N$, where $\{R_i\}$ is an orthonormal basis for $H_a^2\theta\mathcal{B}$.

COROLLARY. *Let H_a^∞ denote $\{f \text{ in } H^\infty \mid f(a)=0\}$. Then there exist $\xi_1,\cdots,\xi_N$ in H_a^∞, and we may take the ξ_i to lie in A, such that every f in H_a^∞ can be written in the form:*

$$f=F\cdot\Phi_0+\sum_{i=1}^{N}t_i\xi_i,\quad F \text{ in } H^\infty,\ t_i \text{ scalars.} \tag{7}$$

In particular, Φ_0, $\xi_1,\cdots,\xi_N$ generate H_a^∞ as ideal in H^∞.

Proof. $R_1,\cdots,R_N$ is an orthonormal basis for $H_a^2\theta\mathcal{B}$. Hence $\int R_i\bar{R}_j\,d\mu'=\delta_i^j$. Since we can approximate each R_i in L^2 by functions f in A with $f(a)=0$, we can construct new functions ξ_i in A, $\xi_i(a)=0$, with $\int\xi_i\bar{R}_j\,d\mu'=\delta_i^j$. Fix f in H_a^∞. Put $t_j=\int f\bar{R}_j\,d\mu'$. Then $f^*=f-\sum_{j=1}^{N}t_j\xi_j$ satisfies $\int f^*\cdot\bar{R}_i\cdot d\mu'=0$, all i, and so f^* is divisible by Φ_0. Hence we get (7).

We now need a result of Gleason (unpublished).

GLEASON'S LEMMA. *Let $\mathfrak{A}$ be any commutative semi-simple Banach algebra with identity, $\mathfrak{M}$ its maximal ideal space, m a point of $\mathfrak{M}$. Assume m is generated, as ideal of $\mathfrak{A}$, by $g_1,\cdots,g_n$. Then there is a neighborhood $\mathfrak{N}$ of m in $\mathfrak{M}$ and a polycylinder $P_r\colon |z_i|<r$, $i=1,\cdots,n$ in C^n, $r>0$, such that the map: $x\rightarrow(g_1(x),\cdots,g_n(x))$ maps $\mathfrak{N}$ homeomorphically on a subset V of P_r, and*

(8) *V is an analytic subvariety of P_r, containing 0, and*

(9) *For each h in $\mathfrak{A}$ there is a function ϕ_h analytic in P_r with*
$h=\phi_h(g_1,\cdots,g_n)$ in $\mathfrak{N}$.

We now apply this Lemma to the case $\mathfrak{A}=H^\infty$, $\mathfrak{M}=S(H^\infty)$ is the maximal ideal space of H^∞. Take for m the homomorphism $\tilde{a}\colon f\rightarrow f(a)$. By (7), H_a^∞ is generated by Φ_0, $\xi_1,\cdots,\xi_N$. Let P_r and V denote the associated polycylinder of C^{N+1} and analytic subvariety of P_r of the preceding Lemma, and let $\mathfrak{N}$ be the associated neighborhood of $\tilde{a}$ in $S(H^\infty)$.

LEMMA 8. *There is some r_0, $0<r_0\leqq r$, such that $P_{r_0}\cap V$ meets each hyperplane $z_1=\lambda$ at most finitely often.*

Proof. Consider the Banach space $H^\infty\oplus C^{N+1}$ which is the direct sum of H^∞ and C^{N+1}. Let T be the linear map[4] of $H^\infty\oplus C^{N+1}$ into H^∞ which sends

$$(F, t_1, \cdots, t_N, t) \rightarrow F \cdot \Phi_0 + \sum_1^N t_i \xi_i + t.$$

By (7), T maps onto H^∞. Also T is one-one. Hence every linear transformation T' of $H^\infty \oplus C^{N+1}$ into H^∞ with $\| T' - T \| < \epsilon$, for some $\epsilon > 0$, maps onto H^∞. For each $r' < r$, put $\mathfrak{N}_{r'}$ equal to the subneighborhood of $\mathfrak{N}$ which the map $\chi \colon x \rightarrow (\Phi_0(x), \xi_1(x), \cdots, \xi_n(x))$ maps on $V \cap P_{r'}$. For x in $\mathfrak{N}_{r'}$ define a linear map T_x by

$$T_x(F, t_1, \cdots, t_N, t) = F \cdot (\Phi_0 - \Phi_0(x)) + \sum_{i=1}^N t_i(\xi_i - \xi_i(x)) + t.$$

Then fix r_0 such that for x in $\mathfrak{N}_{r_0}$ $\| T_x - T \| < \epsilon$, and so T_x maps onto H^∞. If now x is in $\mathfrak{N}_{r_0}$ and f is in H^∞,

$$f = F \cdot (\Phi_0 - \Phi_0(x)) + \sum_{i=1}^N t_i(\xi_i - \xi_i(x)) + t,$$

whence $f \equiv \sum_{i=1}^N t_i(\xi_i - \xi_i(x)) + t$ modulo the ideal generated by $\Phi_0 - \Phi_0(x)$. Hence for each j the functions $1, \xi_j, \cdots, \xi_j^{N+1}$ are linearly dependent modulo that ideal, and so for some polynomial P_j:

$$(10) \qquad P_j(\xi_j) = F_j \cdot (\Phi_0 - \Phi_0(x)) \text{ for some } F_j \text{ in } H^\infty.$$

Hence if two points x and y in $\mathfrak{N}_{r_0}$ have images $\chi(x)$ and $\chi(y)$ in $V \cap P_{r_0}$ in the hyperplane: $z_1 = \lambda$, i. e. if $\Phi_0(x) = \Phi_0(y) = \lambda$, then

$$P_j(\xi_j(y)) = F_j(y)(\Phi_0(y) - \Phi_0(x)) = 0$$

and so $\xi_j(y)$ is one of the finitely many roots of the equation $P_j(z) = 0$. It follows that only finitely many points of $V \cap P_{r_0}$ lie in the hyperplane: $z_1 = \lambda$.

COROLLARY. *If z is a regular point of V sufficiently close to* 0, *then the dimension of V at z is* 1 *or* 0.

LEMMA 9. *In a suitable neighborhood U of* 0 *in C^{N+1} we have*

$$U \cap V = D_1 \cup D_2 \cup \cdots \cup D_k,$$

where for each i there exists a one-one analytic map ψ_i of $|\lambda| < 1$ on D_i, $i = 1, \cdots, k$, and $\psi_i(0) = 0$.

Proof. Let (V) denote the germ of the variety V at 0, and let (V_i), $i = 1, \cdots, k$ be the irreducible branches of (V), so that

---

[4] Such a map is used in the proof of Gleason's Lemma.

$$(V) = (V_1) \cup \cdots \cup (V_k).$$

Because of the preceding Corollary each (V_i) has dimension 1 or 0. We claim each (V_i) has dimension 1.

For else (V) reduces to the single point 0, and thus 0 is an isolated point on V. Since χ is a homeomorphism from $S(H^\infty)$ into C^{N+1} on some neighborhood of $\tilde{a}$, then $\tilde{a}$ is isolated in $S(H^\infty)$. It follows by a well-known result of Silov that there exists e in H^∞ with $e(a) = 0$, $e(m) = 1$ on the rest of $S(H^\infty)$. Hence $\| e \| = 1$ and $e(b) = 1$. Now for f in H^∞

$$f(a) = \int f \cdot K \, d\mu_b, \; f(b) = \int f \, d\mu_b, \; K \text{ in } L^1(\mu_b), \; K \geqq 0.$$

It follows that a and b belong to the same part of $S(H^\infty)$. But this contradicts that $e(a) = 0$, $e(b) = 1$, $\| e \| = 1$, by the discussion preceding Lemma 2. (The argument here was suggested by Andrew Browder).

Thus each (V_i) is a 1-dimensional irreducible germ of variety at 0. The assertion of the Lemma follows from this.

Note. We shall consider V as identified with $\chi^{-1}(V)$ in $S(H^\infty)$. The elements of H^∞ can then be regarded as functions on V, and $\tilde{a}$ as the origin in C^{N+1}.

LEMMA 10. *Fix i, $1 \leqq i \leqq k$. Put $D = D_i$ of the previous lemma. Then there exists f in A with $f(0) = 0$, but f not identically 0 on D.*

Proof. Suppose the contrary. Then f in A and $f(a) = 0$ implies $f \equiv 0$ on D.

Since Φ_0 is in H^2, and $\Phi_0(a) = 0$, we can choose $\{f_n\}$ in A, $f_n(a) = 0$, so that $f_n \to \Phi_0$ in $L^2 = L^2(\mu_b)$. Recall the functions $\xi_1, \cdots, \xi_N$ of the Corollary to Lemma 7. By (7), $f_n = F_n \cdot \Phi_0 + \sum_{i=1}^{N} t_i^{(n)} \xi_i$, where $F_n \in H^\infty$, the $t_i^{(n)}$ are scalars. Let the R_i be as in Lemma 7. Then

$$\int f_n \cdot \bar{R}_i \, d\mu' = t_i^{(n)} \to \int \Phi_0 \cdot \bar{R}_i \, d\mu' = 0.$$

Hence $F_n \cdot \Phi_0 \to \Phi_0$ in L^2 and so $F_n \to 1$ in L^2, since $| \Phi_0 |$ is bounded away from 0. It follows that $F_n(a) \to 1$. Hence for large $n, F_n(a) \neq 0$. As a function on D, F_n is continuous and $\neq 0$ at 0, hence $\neq 0$ in some neighborhood of 0. Also $F_n \cdot \Phi_0$ is in A, since f_n is in A and the ξ_i are in A, and $F_n \cdot \Phi_0(a) = 0$. Hence by supposition, $F_n \cdot \Phi_0 \equiv 0$ on D. It follows that Φ_0 vanishes in some neighborhood of 0 on D. Hence D meets the hyperplane: $z_1 = 0$ infinitely often, and this contradicts Lemma 8. Hence our assertion is correct.

Proof of Theorem 1. As before, fix i and put $D = D_i$, and put $\psi = \psi_i$ where ψ_i is the one-analytic map of $|\lambda| < 1$ on D existing by Lemma 9.

Let A_1 be the collection of functions $f(\psi)$ with f in A. Then A_1 is an algebra of analytic functions on $|\lambda| < 1$. By Lemma 10, there exists f_0 in A with $f_0(0) = 0$ and $f_0 \not\equiv 0$ on D. Then $f_0(\psi)$ belongs to A_1 and without loss of generality we may assume $f_0(\psi) = \lambda^k(1 + c_1\lambda + \cdots)$. Hence $f_0(\psi) = \xi^k$, where ξ is a function analytic and one-one in some neighborhood of 0. A suitable neighborhood of 0 is then mapped by ξ conformally and one-one on a disk $|\xi| < r$.

For each f in A, define F in $|\xi| < r$ by $f(\psi(z)) = F(\xi(z))$. Then F is analytic in $|\xi| < r$. Let A_2 denote the algebra of all these functions F. We do not know whether or not A_2 separates points on $|\xi| < r$. We know that ξ^k is in A_2. By an elementary argument one may now show the following: there exists $\beta = \exp 2\pi i/l$ and r_0, $0 < r_0 < r$, such that in the disk $|\xi| < r_0$ two points x and y are identified by the algebra A_2 if and only if $y = \beta^n \cdot x$ for some n.

It follows that for all F in A_2 $F(w^{1/l})$ is an analytic single valued function of w in the disk $|w| < r_0^l$. Fix w in $|w| < r_0^l$. If f is in A, $f(\psi) = F(\xi)$ for some F in A_2. We define a map $\phi(w)$ of A into the scalars by: $f \to F(w^{1/l})$. Clearly $\phi(w)$ is a homomorphism and so a point of M. Fix w_1 and w_2 and suppose $\phi(w_1) = \phi(w_2)$. It follows that $F(w_1^{1/l}) = F(w_2^{1/l})$, all F in A_2, whence $w_1^{1/l} = \beta^n \cdot w_2^{1/l}$ for some n, and so $w_1 = w_2$. Thus ϕ is a one-one map of $|w| < r_0^l$ into M, with $\phi(0) = a$. Further, for f in A, $f(\phi(w)) = F(w^{1/l})$ is an analytic function of w. It follows that ϕ is continuous as a map from the disk to M. Thus a is the center of an analytic disk.

BROWN UNIVERSITY.

REFERENCES.

[1] J. Wermer, "Dirichlet algebras," *Duke Mathematical Journal*, vol. 27 (1960), pp. pp. 373-382.

[2] A. Gleason, "Function algebras," *Seminars on analytic functions*, Institute for Advanced Study, Princeton (1957), vol. II.

[3] K. Hoffman, "Analytic functions and logmodular Banach algebras," *Acta Mathematica*, vol. 108 (1962).

[4] R. Arens and I. M. Singer, "Function values as boundary integrals," *Proceedings of the American Mathematical Society*, vol. 5 (1954), pp. 735-745.

SPHERICAL FUNCTIONS OVER $\mathfrak{P}$-ADIC FIELDS, II.*

By F. I. MAUTNER.

1. Introduction and preliminaries. This is a continuation of Part I (see Mautner [1]) and the same notation and terminology will be used as in Part I. We shall study certain classes of functions on the projective group $G = PGL(2, \Omega)$ over a $\mathfrak{P}$-adic field Ω, i. e. a field which is complete under a given discrete real valued valuation $|\ |$ and which has a finite residue class field $\mathfrak{O}/\mathfrak{P}$ with $q = \mathfrak{N}\mathfrak{P}$ elements (see § 1 of Part I).

As in Part I let K be the subgroup of G consisting of $\mathfrak{P}$-adic integral linear fractional transformations. We know that K is a maximal compact subgroup of G which is also open in G (see §§ 1 and 2 of Part I). Now let u be an irreducible continuous unitary representation of K acting in a finite dimensional vector space $\mathfrak{v}$ over the complex numbers.

Definition 1.1. Let S_u be the class of all functions $\mathfrak{f}(g)$ on G whose values are linear transformations of $\mathfrak{v}$ and which satisfy

(1.1) $\qquad \mathfrak{f}(kgk') = u(k)\mathfrak{f}(g)u(k')$ for all $g \in G, k$ and $k' \in K$.

By S_u^0 we denote the class of functions in S_u which are of compact support on G. Note that $\mathfrak{f}(g) \in S_u$ implies that $\mathfrak{f}(g)$ is continuous, since G/K is discrete and $u(k)$ continuous on K. One sees that S_u^0 is an algebra with convolution on G as product.

We shall study the algebra S_u^0 for certain irreducible representations u of K (the principal series mod $\mathfrak{P}^m$, see § 3 below). For these $u = u_\chi$ which are naturally parameterized by primitive congruence characters χ, we shall describe in detail a subalgebra $\mathfrak{S}_u^0$ of S_u^0 on which the principal series $M(g, \alpha)$ of representations of G (introduced in Part I) will turn out to be one-one. It will be shown that the Fourier transform on G maps $\mathfrak{S}_u^0$ isomorphically onto the algebra of all Fourier polynomials and the Plancherel formula will be obtained for the corresponding subspace $\mathfrak{S}_e^2$ of $L^2(G)$ (see §§ 6 and 7). For this purpose the principal series $M(g, \alpha)$ has to be studied in some detail in § 3 below and some matrix coefficients $M_{\mathfrak{n}}(g, \alpha)$ of $M(g, \alpha)$ have to be computed (in § 4). This allows one to compute the Fourier trans-

---

Received May 27, 1963.

* Most of this was written several years ago. Some details have hecently been added under NSF Grant G-230011.

forms on G with respect to the principal series $M(g,\alpha)$ quite explicitly in terms of ordinary Fourier series for functions in S_e^2 (§ 5) which leads to the Plancherel formula for $\mathfrak{S}_e^2$. However the full Plancherel formula for $L^2(G)$ remains unknown.[1a] In fact it will be seen in § 8 that the principal series $M(g,\alpha)$ is not one-one on functions of compact support, not even on S_e^0 (or S_u^0). The question thus arises what the other irreducible unitary representations of G are.

Some of them will be obtained in § 9 where it will be shown that there exist discrete series of irreducible unitary representations of G. Their matrix-coefficients (in suitable orthonormal bases) will be seen to be functions of compact support on G!

These results have various arithmetic consequences. If one lets the algebra S_u^0 act on suitable function spaces defined on spaces of idele-classes of quaternion algebras one obtains natural generalizations of Hecke's operators. In particular if one restricts oneself to $\mathfrak{S}_u^0$ for $u = u_\chi$ one can locate the zeros of the characteristic polynomials $Z(s,\chi)$ of these generalized Hecke operators. Indeed one obtains in this manner for each primitive congruence character χ a polynomial $Z(s,\chi)$ in q^s which satisfies $Z(1-s,\chi) = Z(s,\bar{\chi})$ and has an analogue of an Euler product. For $\chi = 1$ (and suitable quaternion algebras) $Z(s,1)$ is a congruence-zeta function of Artin, Hasse and Weil for certain curves mod q. However if χ is a non-real character no such interpretation is known, but the results of the present paper (especially § 7) imply that $Z(s,\chi)$ has all its zeros on $Rs = \frac{1}{2}$. For real χ I have not been able to decide whether $Z(s,\chi)$ may also have real zeros. This seems to be as difficult as Ramanujan's conjecture on the order of magnitude of cusp forms and is closely related to it. The coefficients of $Z(s,\chi)$ can be determined explicitly in terms of class numbers of certain quadratic fields or equivalently in terms of Siegel's representation functions of certain ternary quadratic forms. All this will be discussed in detail elsewhere.

2. The algebras S_e^0 and S_e^1.

Let us introduce an arbitrary orthonormal basis in the r-dimensional representation space $\mathfrak{v}$ of the given unitary representation u of K and denote by $u_{ji}(k)$ the matrix-coefficients of $u(k)$ with respect to this basis. Then the function[2]

$$e(k) = r u_{11}(k) \tag{2.1}$$

[1a] *Added in proof*: In the meantime several articles by I. M. Gelfand and M. I. Graev have appeared on this subject. See *Doklady Akademii Nauk SSSR* 146, pp. 757-760, 147, pp. 529-532, 151, pp. 262-264 and *Uspehi Matematičeskih Nauk* XVIII, pp. 29-99.

[2] Assuming that the Haar measure is normalized so that the total measure of K equals one.

is for fixed $i = 1, 2, \cdots, r$, a self-adjoint minimal idempotent on K,

$$(2.2) \qquad e(k) = \bar{e}(k^{-1}) = \int_K e(k\mathfrak{k}^{-1}) e(\mathfrak{k})\, d\mathfrak{k},$$

as follows at once from the well known orthogonality relations on the compact group K.

Definition 2.1. Let S_e^p denote the set of all (equivalence classes of) complex valued Haar-Lebesgue measurable function $f(g)$ on G which satisfy $\int f(g)|^p\, dg < \infty$ and

$$(2.3) \qquad f(g) = \int_K e(k) f(k^{-1}g)\, dk = \int_K f(gk^{-1}) e(k)\, dk.$$

One verifies at once that S_e^1 is a closed subalgebra of the normed algebra $L^1(G)$.

By S_e^0 we denote the set of all $f(g) \in S_e^1$ which are of compact support on G. One notes that S_e^0 is a subalgebra of the group-algebra of all continuous functions of compact support on G.

Let $\mathfrak{f}(g) \in S_u$. For each $g \in G$ the value of $\mathfrak{f}(g)$ is by definition a linear transformation of $\mathfrak{v}$. With respect to the above orthonormal basis in $\mathfrak{v}$ let $f_{ji}(g)$ denote the matrix-coefficients of this linear transformation.

LEMMA 2.1. *The mapping*

$$(2.4) \qquad \mathfrak{f}(g) \to r f_{ii}(g)$$

is for each fixed $i = 1, 2, \cdots, r$ *a *-isomorphism of* S_u^0 *onto* S_e^0.

Proof. Let us note first that $\mathfrak{f}(g) \in S_u^0$ implies $f_{ii}(g) \in S_e^0$. Indeed using the orthogonality relations on K we have

$$\int_K f_{ii}(gk^{-1}) e(k)\, dk = \int_K \sum_j f_{ij}(g) u_{ji}(k^{-1}) r u_{ii}(k)\, dk$$

$$= \int_K f_{ii}(g) \bar{u}_{ii}(k) r u_{ii}(k)\, dk = f_{ii}(g)$$

by the definition of $e(k)$ and S_u. Similarly

$$\int_K e(k) f_{ii}(k^{-1}g)\, dk = f_{ii}(g).$$

Clearly $\mathfrak{f}(g)$ of compact support implies that $f_{ji}(g)$ is of compact support, so that $f_{ii}(g) \in S_e^0$. The mapping (2.4) is clearly linear. To prove that it is one-one let us note that $\mathfrak{f}(g) \in S_u$ implies

$$(2.5) \qquad f_{lj}(g) = \int_K \int_K r u_{li}(k) f_{ii}(k^{-1} g \mathfrak{k}^{-1}) r u_{ij}(\mathfrak{k})\, dk\, d\mathfrak{k}$$

which is verified at once, using the orthogonality relations on K and (1.1) above. Hence $f_{ii}(g) = 0$ for all $g \in G$ implies $f_{lj}(g) = 0$ for all $g \in G$ and $l, j = 1, \cdots, r$. Thus (2.4) is one-one. We now prove that (2.4) preserves convolution. Let $\mathfrak{a}(g) \in S_u{}^0$ and $\mathfrak{b}(g) \in S_u{}^0$ and denote by $a_{ij}(g)$ and $b_{ij}(g)$ their matrix-coefficients. Then the i-th diagonal matrix-coefficient of $\int \mathfrak{a}(gh^{-1}) \mathfrak{b}(h)\, dh$ equals $\int \sum_l a_{il}(gh^{-1}) b_{li}(h)\, dh$. By (2.5) this equals

$$r^2 \sum_{l=1}^{r} \int_G \int_K \int_K a_{ii}(gh^{-1}k^{-1}) u_{il}(k) u_{li}(\mathfrak{k}) b_{ii}(\mathfrak{k}^{-1}h)\, dk\, d\mathfrak{k}\, dh.$$

Let $u'_{ij}(g)$ be the extension of $u_{ij}(g)$ to G defined by putting

$$u'_{ij}(g) = \begin{cases} u_{ij}(g) & \text{if } g \in K \\ 0 & \text{if } g \in K \end{cases}$$

so that $u'_{ij}(g)$ is a continuous function on G of compact support. Then the integrals over K can be replaced by integrals over G in the last expression. Hence, if $*$ denotes convolution on G we have

$$\sum_{l=1}^{r} a_{il} * b_{li} = \sum_{l=1}^{r} r^2 a_{ii} * u'_{il} * u'_{li} * b_{ii} = r^2 a_{ii} * u'_{ii} * b_{ii}$$

by the orthogonality relations on K. But $r u'_{ii} * b_{ii} = b_{ii}$ since $\mathfrak{b}(g) \in S_u{}^0$, whence $b_{ii}(g) \in S_e{}^0$ as noted above. This proves

$$(2.6) \qquad \sum_{l=1}^{r} a_{il} * b_{li} = r a_{ii} * b_{ii}.$$

It remains to prove that the mapping (2.4) is onto $S_e{}^0$. Let $f(g) \in S_e{}^0$. Define $f_{lj}(g)$ by $f_{lj} = r u'_{li} * r^{-1} f * r u'_{ij}$. For each $g \in G$ let $\mathfrak{f}(g)$ be the linear transformation of $\mathfrak{v}$ whose matrix is $f_{lj}(g)$ with respect to the basis introduced in $\mathfrak{v}$ at the beginning of this section. One verifies at once that $\mathfrak{f}(g) \in S_u{}^0$ and that the image of $\mathfrak{f}(g)$ under (2.4) equals $f(g)$.

Let us now consider the double coset of $Ky^\nu K$ for fixed $\nu > 0$. The group $K \times K$ acts transitively on $Ky^\nu K$ by transforming $g \in Ky^\nu K$ into $kg\mathfrak{k}^{-1}$ with $k, \mathfrak{k} \in K$. Let us find the subgroup which leaves the point y^ν fixed: $ky^\nu \mathfrak{k}^{-1} = y^\nu$, or in terms of matrices

$$\begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} \begin{pmatrix} \tau^\nu & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \tau^\nu & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathfrak{k}_{11} & \mathfrak{k}_{12} \\ \mathfrak{k}_{21} & \mathfrak{k}_{22} \end{pmatrix}$$

i. e.

$$(2.7) \qquad \begin{pmatrix} k_{11}\tau^\nu & k_{12} \\ k_{21}\tau^\nu & k_{22} \end{pmatrix} = \begin{pmatrix} \tau^\nu \mathfrak{k}_{11} & \tau^\nu \mathfrak{k}_{12} \\ \mathfrak{k}_{21} & \mathfrak{k}_{22} \end{pmatrix}.$$

Now all k_{ij} and $\mathfrak{k}_{ij}$ are elements of $\mathfrak{O}$, so that we must have $k_{12} \in \mathfrak{P}^{\nu}$ and $\mathfrak{k}_{21} \in \mathfrak{P}^{\nu}$. Let $\mathfrak{K}_{\nu}$ be the subgroup of $\mathfrak{K} = GL(2, \mathfrak{O})$ with $k_{12} \in \mathfrak{P}^{\nu}$ and $\mathfrak{K}_{\nu}{}^{t}$ the group of transposed matrices. The map

$$(2.8) \qquad (k_{ij}) \to \begin{pmatrix} \tau^{-\nu} & 0 \\ 0 & 1 \end{pmatrix} (k_{ij}) \begin{pmatrix} \tau^{\nu} & 0 \\ 0 & 1 \end{pmatrix}$$

is clearly an isomorphism of $\mathfrak{K}_{\nu}$ onto $\mathfrak{K}_{\nu}{}^{t}$. Let us apply the natural homomorphism of $\mathfrak{G} = GL(2, \Omega)$ onto $G = PGL(2, \Omega)$ to this situation. Letting K_{ν} and $K_{\nu}{}^{t}$ be the images of $\mathfrak{K}_{\nu}$ and $\mathfrak{K}_{\nu}{}^{t}$, we obtain from (2.8) a well defined isomorphism of K_{ν} onto $K_{\nu}{}^{t}$ which we denoted by (ν). And

$$(2.9) \qquad k y^{\nu} \mathfrak{k}^{-1} = y^{\nu} \text{ if and only if } k \in K_{\nu} \text{ and } \mathfrak{k} = k^{(\nu)}.$$

Now let u be a representation of K in an r-dimensional vector space $\mathfrak{v}$. We wish to find all $r \times r$ matrices X satisfying

$$u(k) X u(\mathfrak{k}) = X \text{ whenever } k \in K_{\nu} \text{ and } \mathfrak{k}^{-1} = k^{(\nu)},$$

i. e.

$$(2.10) \qquad u(k) X = X u(k^{(\nu)}) \text{ for all } k \in K_{\nu}.$$

Let us write $u_{\nu}(k) = u(k^{(\nu)})$ for $k \in K_{\nu}$, then u_{ν} is a representation of K_{ν} and (2.10) can be rewritten

$$(2.11) \qquad u(k) X = X u_{\nu}(k) \text{ for all } k \in K_{\nu}.$$

Thus we are looking for all matrices X which "intertwine" the two representations u, u_{ν} of the compact group K_{ν} and have proved

LEMMA 2.2. *The elements* $\mathfrak{f}$ *of* S_u *which vanish outside the double coset* $K y^{\nu} K$ *are in natural one-one correspondence with the matrices* X *which satisfy* (2.11).

Let us make

Definition 2.2. We denote by $S_u(\nu)$ the set of all $\mathfrak{f} \in S_u$ which vanish outside the double coset $K y^{\nu} K$, similarly by $S_e(\nu)$ the space of those $f \in S_e$ which vanish outside $K y^{\nu} K$.

Then we can restate Lemma 2.2 more precisely: Let X denote the restriction of $\mathfrak{f}$ to y^{ν}. The map $\mathfrak{f} \to X$ maps $S_u(\nu)$ one-one linearly onto the space of all X satisfying (2.11).

3. The principal series of representations of G. As in Part I, let $M(g, \alpha)$ denote again the principal series of representations of G defined as induced representations

$$(3.1) \qquad \operatorname*{ind}_{AB\uparrow G} \alpha$$

where α is a character of A extended to a one-dimensional representation of AB by putting $\alpha(ab) = \alpha(a)$. (Cf. § 3 of Part I, also Mackey [1] for the definition and basic properties of induced representations.) The representation space of $M(g, \alpha)$ is a space of functions $\Psi(g)$ on G satisfying $\Psi(abg) = \alpha(a)\Psi(g)$. We note that $\Psi(g)$ is determined by its restriction to K, because $G = ABK$.

Now let us consider the restriction of $M(g, \alpha)$ to K. This can be described by

$$[M(k, \alpha)\Psi](\mathfrak{k}) = \Psi(\mathfrak{k}k) \text{ and } \Psi(abk) = \alpha(a)\Psi(k)$$

for all $k, \mathfrak{k} \in K$, $a \in A \cap K$, $b \in B \cap K$, so that $\alpha(a) = \alpha_1(a)$ where α_1 is the restriction of α to $A \cap K$. This proves

$$(3.2) \qquad \text{restriction to } K \text{ of } \operatorname*{ind}_{AB\uparrow G} \alpha = \operatorname*{ind}_{(K \cap AB)\uparrow K} \alpha_1$$

where α_1 = restriction of α to $A \cap K$, considered as one-dimensional representation of $K \cap AB$.

Now let u be an irreducible continuous unitary representation of K. By the Frobenius reciprocity theorem u is contained in the induced representation (3.2) above as often as the restriction of u to $K \cap AB$, which we denote by $u(K \cap AB)$, contains $\alpha_1(K \cap AB)$.

Now we consider the following special kind of representation u. For fixed m, let $\bar{K} = \bar{K}^{(m)}$ denote[8] the projective group $PGL(2, q^m)$ over the ring $\mathfrak{O}/\mathfrak{P}^m$. There exists a natural homomorphism of K onto $\bar{K}$, namely reduction modulo $\mathfrak{P}^m$. Let $\bar{A}$ and $\bar{B}$ be the images of A, respectively B, under this homomorphism, so that $\bar{A}$ is isomorphic to the multiplicative group of invertible elements of $\mathfrak{O}/\mathfrak{P}^m$, $\bar{B}$ to the additive group of $\mathfrak{O}/\mathfrak{P}^m$. Let χ be a character of $\bar{A}$, which we extend to a one-dimensional representation of $\bar{A}\bar{B}$ by putting $\chi(\bar{a}\bar{b}) = \chi(\bar{a})$. Now let $\bar{u}$ be the induced representation

$$(3.3) \qquad \bar{u} = \operatorname*{ind}_{\bar{A}\bar{B}\uparrow\bar{K}} \chi.$$

Let us apply again the Frobenius reciprocity theorem to this: We see that if $\bar{u}$ is irreducible, then the restriction $\bar{u}(\bar{A}\bar{B})$ of $\bar{u}$ to $\bar{A}\bar{B}$ contains χ exactly once:

$$(3.4) \qquad \text{multiplicity of } \chi(AB) \text{ in } \bar{u}(\bar{A}\bar{B}) = 1,$$

---

[8] Throughout this section $\bar{K}$, $\bar{A}$, $\bar{B}$, $\bar{u}, \cdots$ do not denote complex conjugates; elsewhere $\bar{z}$ denotes the complex conjugate of any complex number z.

provided $\bar{u}$ is an irreducible representation of $\bar{K}$ given by (3.3).

If we combine the homomorphism $K \to \bar{K}$ with the representation $\bar{u} = \bar{u}_\chi$ of $\bar{K}$ defined by (3.3) we obtain a unitary representation of K which we denote by u_χ:

$$u_\chi(k) = \bar{u}_\chi(\bar{k}) \text{ where } \bar{k} = K^{(m)}k \tag{3.5}$$

if k is mapped into $\bar{k}$ under the homomorphism $K \to \bar{K}$ and $K^{(m)}$ denotes the subgroup of those elements k of K which are congruent to the identity modulo $\mathfrak{P}^m$, i.e. k is given by a matrix (k_{ij}) such that

$$k_{ij} \equiv \delta_{ij} \pmod{\mathfrak{P}^m}. \tag{3.6}$$

Now let us assume that α_1 is identically one on $1 + \mathfrak{P}^m$ so that we can identify α_1 with a character of $\bar{A}$ (= image of A under the homomorphism $K \to \bar{K}$). Let χ also be a character of $\bar{A}$. We wish to find the multiplicity of $u = u_\chi$ (as defined by (3.5) above) in the representation

$$\operatorname*{ind}_{(K \cap AB)\uparrow K} \alpha_1. \tag{3.7}$$

By the Frobenius reciprocity theorem this multiplicity equals the multiplicity of α_1 (considered as a character of $AB \cap K$) in $u(AB \cap K)$ [= restriction of u to $AB \cap K$]. By taking the homomorphism $K \to \bar{K}$ the latter equals the multiplicity of α_1 (considered as a character of $\bar{A}\bar{B}$) in the restriction $\bar{u}(\bar{A}\bar{B})$ of $\bar{u}$ to $\bar{A}\bar{B}$ where $\bar{u}$ is defined by (3.3). Now let us assume that $\bar{u} = \bar{u}_\chi = \operatorname{ind}\chi$ is an irreducible representation of $\bar{K}$ and let us decompose its restriction to $\bar{A}\bar{B}$ into inequivalent irreducible representations ρ of $\bar{A}\bar{B}$ occurring with multiplicity $m_{\rho\chi}$:

$$\bar{u}_\chi(\bar{A}\bar{B}) = \sum_\rho{}^{\oplus} m_{\rho\chi}\, \rho. \tag{3.8}$$

Applying again the Frobenius reciprocity theorem we see that

$$\operatorname*{ind}_{\bar{A}\bar{B}\uparrow\bar{K}} \rho \text{ contains } \bar{u}_\chi \text{ exactly } m_{\rho\chi} \text{ times.} \tag{3.9}$$

In particular let $\rho(\bar{a}\bar{b}) = \alpha_1(\bar{a})$ and we see that

$$\begin{aligned} m_{\alpha_1,\chi} &= \text{multiplicity of } \alpha_1(\bar{A}\bar{B}) \text{ in } \bar{u}_\chi(\bar{A}\bar{B}) \\ &= \text{multiplicity of } \bar{u}_\chi \text{ in } \operatorname*{ind}_{\bar{A}\cdot\bar{B}\uparrow\bar{K}} \alpha_1 = \bar{u}_{\alpha_1}. \end{aligned} \tag{3.10}$$

Hence if α_1 is such that $\bar{u}_{\alpha_1}$ is also irreducible, then (3.10) equals 1 or 0 according as $\bar{u}_\chi$ and $\bar{u}_{\alpha_1}$ are equivalent representations of $\bar{K}$ or not.

We now have to decide when u_χ is irreducible. This is of course not new.

LEMMA 3.1. *Let* χ *be a non-real primitive congruence character* mod $\mathfrak{P}^m$, *i. e. let* χ *be a non-real character of* $\mathfrak{U}$ *and* m *the smallest integer such that* χ *is one on* $1+\mathfrak{P}^m$. *Then the principal series* mod $\mathfrak{P}^m$ *as defined by* (3.3) *is an irreducible representation of* $PGL(2,\mathfrak{O}/\mathfrak{P}^m)=\bar{K}$.

Proof. The representation $\bar{u}=\bar{u}_\chi$ is defined (see (3.3) above) to be the induced representation $\operatorname{ind}\chi$. Hence the commuting algebra is given by the function $f(k)$ defined on $\bar{K}$ and satisfying $f(\bar{a}\bar{b}\bar{k})=\chi(\bar{a})f(\bar{k})=f(\bar{k}\bar{a}\bar{b})$. So we have to consider the double cosets $\overline{AB}\,\bar{k}\,\overline{AB}$ of AB in $\bar{K}$. They are easily seen to be $\overline{AB}$, $\overline{AB}\,\bar{w}\,\overline{AB}$ and $\overline{AB}\begin{pmatrix}1 & 0\\ \tau^l & 1\end{pmatrix}\overline{AB}$ with $l=1,\cdots,m-1$. Now

$$\begin{pmatrix}a & b\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \tau^l & 1\end{pmatrix}=\begin{pmatrix}a+b\tau^l & b\\ \tau^l & 1\end{pmatrix}$$

and

$$\begin{pmatrix}1 & 0\\ \tau^l & 1\end{pmatrix}\begin{pmatrix}a' & b'\\ 0 & 1\end{pmatrix}=\begin{pmatrix}a' & b'\\ \tau^l a' & \tau^l b'+1\end{pmatrix}$$

give equal elements of $\bar{K}$ if and only if

$$\left.\begin{array}{l} a'\equiv\epsilon(a+b\tau^l),\, b'\equiv\epsilon b\\ \tau^l a'\equiv\epsilon\tau^l,\, 1+\tau^l b'\equiv\epsilon\end{array}\right\}\pmod{\mathfrak{P}^m}$$

for some $\epsilon\in\mathfrak{U}$. Thus $1+\tau^l\epsilon b\equiv\epsilon$, so that $a'\equiv\epsilon(a+1)-1\pmod{\mathfrak{P}^m}$ and $a'\tau^l\equiv\epsilon\tau^l\equiv\epsilon\tau^l(a+1)-\tau^l$ gives $\tau^l(\epsilon a-1)\equiv 0\pmod{\mathfrak{P}^m}$. We can certainly find elements a and ϵ of $\mathfrak{U}$ such that $\tau^l(\epsilon a-1)\equiv 0\pmod{\mathfrak{P}^m}$ then define a' by $a'\equiv\epsilon(a+1)-1\pmod{\mathfrak{P}^m}$ and have $\chi(a)\neq\chi(a')$ since χ is a primitive character mod $\mathfrak{P}^m$. Hence any function $f(\bar{k})$ satisfying $f(\bar{a}\bar{b}\bar{k})=\chi(\bar{a})f(\bar{k})=f(\bar{k}\bar{a}\bar{b})$ must vanish on $\bar{A}\bar{B}\begin{pmatrix}1 & 0\\ \tau^l & 1\end{pmatrix}\bar{A}\bar{B}$.

It remains to consider $\overline{AB}\bar{w}\overline{AB}$. Since $\bar{a}\bar{w}\bar{a}=\bar{a}$ in $\bar{K}$ we must have $f(\overline{aw}\,\bar{a})=f(\bar{w})$ hence $\chi(\bar{a}^2)f(\bar{w})=f(\bar{w})$. By hypothesis χ is not real, i.e. $\chi(\bar{a})^2\neq 1$ for some $\bar{a}$, hence $f(\bar{w})=0$. This proves that the commuting algebra of $\bar{u}_\chi$ consists of functions vanishing outside $\bar{A}\bar{B}$ and clearly $f(\bar{a}\bar{b})=c\chi(\bar{a})$ for some complex number c, hence consists of the scalar multiples of the identity operator only. This proves Lemma 3.1.

If we combine Lemma 3.1 with the result (proved above) that (3.10) equals 1 or 0 according as u_χ and u_{α_1} are equivalent or not provided they are irreducible, we obtain

THEOREM 3.1. *Let* α *be a character of* A *and* α_1 *its restriction to* $A\cap K$, *so that* α_1 *can be considered to be a character of* $\mathfrak{U}$. *Assume that* $\alpha_1(\eta)$ *is not real for some* $\eta\in\mathfrak{U}$. *Let* m *be the smallest integer such that* α_1 *is* *identically one on* $1+\mathfrak{P}^m$. *Let* χ *be also a primitive non-real character of* $\mathfrak{U}/(1+\mathfrak{P}^m)$ *and form the unitary representation* $u=u_\chi$ *as defined by* (3.5). *Then* u_χ *occurs in the restriction of the principal series* $M(g,\alpha)$ *to* K *either exactly once or not at all according as* $\bar{u}_\chi$ *and* $\bar{u}_{\alpha_1}$ *are equivalent representations of* $\bar{K}$ *or not. Here* $\bar{K}=PGL(2,\mathfrak{O}/\mathfrak{P}^m)$.

One can write down quite explicitly and directly a subspace which transforms according to this u under $M(k,\alpha)$. Indeed the representation space of $M(k,\alpha)$ can be taken to consist of functions $\Psi(k)$ satisfying $\Psi(\mathfrak{a}\mathfrak{b}k)=\alpha(\mathfrak{a})\Psi(k)$ for $k\in K$, $\mathfrak{a}\mathfrak{b}\in AB\cap K$, as we have seen above. Let $K^{(m)}$ be the subgroup of K whose elements are given by $\mathfrak{P}$-adic integral matrices $\equiv I \pmod{\mathfrak{P}^m}$; then we must have $M(k',\alpha)\Psi=\Psi$ for all $k'\in K^{(m)}$, i.e.

$$\Psi(kk')=\Psi(k)\text{ for }k\in K,\ k'\in K^{(m)} \tag{3.11}$$

so that Ψ must be a function on $\bar{K}=K/K^{(m)}$. Now α_1 is by hypothesis a character mod $\mathfrak{P}^m$, i.e. $\alpha_1(1+\mathfrak{P}^m)=1$, so that $\alpha_1(a)=\alpha(\bar{a})$. Thus (3.11) implies that we are dealing with all functions on $\bar{K}$ satisfying $\Psi(\bar{\mathfrak{a}}\bar{\mathfrak{b}}\bar{k})=\alpha_1(\bar{a})\Psi(\bar{k})$ and $[M(\bar{k}',\bar{\alpha}_1)\Psi](\bar{k})=\Psi(\bar{k}\bar{k}')$. This is precisely the definition of the induced representation $\underset{\bar{A}\cdot\bar{B}\uparrow\bar{K}}{\operatorname{ind}}\bar{\alpha}_1$.

This proves

LEMMA 3.2. *The subspace which transforms according to* u_{α_1} *under* $M(k,\alpha)$ *for* $k\in K$ *consists exactly of those functions* Ψ *which satisfy* $M(k',\bar{\alpha})\Psi=\Psi$ *for* $k'\in K^{(m)}$. *Since it is equivalently given by* (3.11) *it has dimension* q^m+q^{m-1}.

We shall use the following expression for $M(g,\alpha)$. The factorspace G/AB is homeomorphic to the projective line over Ω. In every coset except that corresponding to the point at infinity there exists an element $\mathfrak{x}$ of G which is given by a matrix $\begin{pmatrix}1 & 0\\ x & 1\end{pmatrix}$. Consider any function $\Psi(g)$ in the representation space of $M(g,\alpha)$ and its transform Ψ' by $M(g,\alpha)$, i.e. $\Psi'=M(g,\alpha)\Psi$, which means by definition $\Psi'(\mathfrak{g})=\Psi(\mathfrak{g}g)$ for all $\mathfrak{g}$ and $g\in G$. Let $\mathfrak{g}=\mathfrak{x}$ then $\mathfrak{g}g=\mathfrak{x}g=\mathfrak{a}\mathfrak{b}\mathfrak{x}_1$ where $\mathfrak{x}_1$ is given by a matrix $\begin{pmatrix}1 & 0\\ x_1 & 1\end{pmatrix}$, provided $\mathfrak{x}g$ does not belong to the coset mod $A\cdot B$ corresponding to the point at infinity. We obtain $\Psi'(\mathfrak{x})=\Psi(\mathfrak{x}g)=\Psi(\mathfrak{a}\mathfrak{b}\mathfrak{x}_1)=\alpha(\mathfrak{a})\Psi(\mathfrak{x}_1)$ where $\mathfrak{a}$ is given by a diagonal matrix $\begin{pmatrix}a_1 & 0\\ 0 & a_2\end{pmatrix}$ and $\mathfrak{b}$ by a matrix $\begin{pmatrix}1 & b\\ 0 & 1\end{pmatrix}$. Let us consider the matrix equation corresponding to $\mathfrak{x}g=\mathfrak{a}\mathfrak{b}\mathfrak{x}_1$:

$$\begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ x_1 & 1 \end{pmatrix}$$

i. e.

$$(3.12) \qquad \begin{pmatrix} g_{11} & g_{12} \\ xg_{11} + g_{21} & xg_{12} + g_{22} \end{pmatrix} = \begin{pmatrix} a_1 + a_1 b x_1 & a_1 b \\ a_2 x_1 & a_2 \end{pmatrix}$$

Thus $a_2 = xg_{12} + g_{22}$ and $x_1 = a_2^{-1}(g_{11}x + g_{21})$, hence

$$(3.13) \qquad x_1 = \frac{g_{11}x + g_{21}}{g_{12}x + g_{22}}$$

if g is given by the matrix (g_{ij}). Moreover $a_1(1 + bx_1) = g_{11}$ and $a_1 b = g_{12}$ imply

$$(3.14) \qquad a_1 = \frac{\det(g_{ij})}{g_{12}x + g_{22}} \text{ provided } g_{12}x + g_{22} \neq 0.$$

since the map $\begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} \to \mathfrak{x}$ is one-one, we can write $\Psi(\mathfrak{x}) = \Psi(x)$ and obtain $\Psi'(x) = \alpha(a)\Psi(g_{11}x + g_{21}/g_{12} + g_{22})$. Now α is a character of A and $A \cong \Omega^{\times}$. Thus when a is given by a diagonal matrix $\begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$ we shall have $\alpha(a) = \alpha(a_1/a_2)$. Therefore we obtain

$$(3.15) \qquad [M(g,\alpha)\Psi](x) = \alpha[(g_{12}x + g_{22})^{-2}\det(g_{ij})]\Psi\left(\frac{g_{11}x + g_{21}}{g_{12}x + g_{22}}\right),$$

provided $g_{12}x + g_{22} \neq 0$. As g is an element of $G = PGL(2,\Omega)$, the matrix (g_{ij}) can be multiplied by any $c \in \Omega^{\times}$; but we note that the right side of (3.15) does not change if (g_{ij}) is replaced by (cg_{ij}).

Let $w \in G$ be given by $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, i. e. $wx = -x^{-1}$ and consider $M(w,\alpha)$ as an operator on $L^2(\Omega, dx)$. Let us write

$$(3.16) \qquad a = \tau^n \eta \text{ with } n = \operatorname{ord} a,\ \eta \in \mathfrak{U} \text{ and } \alpha(a) = q^{n(s-1)}\alpha_1(\eta)$$

where s is an arbitrary complex number and α_1 = restriction of α to $\mathfrak{U}$ is a character of $\mathfrak{U}$.

LEMMA 3.3. *$M(w,\alpha)$ is a unitary operator of $L^2(\Omega, dx)$ if and only if $\sigma = Rs = \frac{1}{2}$.*

Proof. By (3.15) we have

$$[M(w,\alpha)\Psi](x) = \alpha(x)^{-2}\Psi(-x^{-1})$$

so that[4]

---

[4] As in Part I the symbol $|\ |_{\infty}$ denotes the ordinary absolute value of a complex number throughout, to distinguish it from the normalized $\mathfrak{P}$-adic absolute value $|\ |$.

$$\| M(w,\alpha)\Psi \|^2 = \int | \alpha(x)^{-2}\Psi(-x^{-1}) |_\infty^2 \, dx = \int | \alpha(-x)^2 \Psi(x) |_\infty^2 \, | x |^{-2} \, dx.$$

Hence

$$\| M(w,\alpha)\Psi \| = \| \Psi \| \text{ for all } \Psi \in L^2(\Omega, dx)$$

if and only if $| \alpha(x) |_\infty^2 = | x |$ for all $x \in \Omega^\times$, i.e. if and only if $| \alpha(\tau^n) |_\infty^2 = | \tau^n |$, i.e. $q^{2n(\sigma-1)} = q^{-n}$, i.e. $\sigma = \frac{1}{2}$.

From this we obtain at once

LEMMA 3.4. *$M(g,\alpha)$ is a unitary operator of $L^2(\Omega, dx)$ for all $g \in G$ if and only if $\sigma = Rs = \frac{1}{2}$.*

Proof. Let us note that G is generated by w, A and B^t. If $g \in B^t$ then $M(g,\alpha)$ is unitary for all α. It remains to observe that $M(g,\alpha)$ is unitary for $g \in A$ if $\sigma = \frac{1}{2}$. Indeed if g is given by $\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$ and $s = \frac{1}{2} + it$, with t real, then

$$[M(g,\alpha)\Psi](x) = \alpha(a)\Psi(ax) = q^{n(it-\frac{1}{2})}\alpha_1(\eta)\Psi(ax)$$

so that

$$\begin{aligned} \| M(g,\alpha)\Psi \|^2 &= \int q^{-n} | \Psi(ax) |^2 \, dx = \int q^{-n} | \Psi(x) |_\infty^2 \cdot | a |^{-1} \, dx \\ &= \int | \Psi(x) |_\infty^2 \, dx. \end{aligned}$$

THEOREM 3.2. *Let $\alpha(a)$ be given by (3.16) with $Rs = \frac{1}{2}$. Then $g \to M(g,\alpha)$ is an irreducible unitary representation of G with Hilbert space $L^2(\Omega, dx)$. In fact the restriction to AB is already irreducible.*

Proof. By Lemma 3.3 $M(g,\alpha)$ is unitary. The irreducibility under AB is proved in the same way as in the case of the real and complex numbers. One observes that any bounded operator which commutes with all translations $\Psi(x) \to \Psi(x+b)$ is in the algebra generated by such translations, i.e. after Fourier transform on the additive group of Ω is multiplication by a bounded measurable function. But $\Omega^\times$ is transitive on the non-constant additive characters of Ω, so that if C is to commute also with all $\Psi(x) \to \alpha(a)\Psi(ax)$, it must be a scalar multiple of the identity: $C = cI$.

The irreducibility of the principal series (for almost all α) was proved in the case of the classical groups over $\mathfrak{P}$-adic fields by Bruhat [1].

Let χ be a congruence character mod $\mathfrak{P}^m$ as above, then the above implies that the representation $\tilde{u}_\chi$ defined by (3.3) can be written as an explicit multiplier representation analogous to (3.15) as follows. Let x vary over the projective line over $\mathfrak{O}/\mathfrak{P}^m$ and $\Psi(x)$ an arbitrary complex valued function of x, then for $k \in \bar{K} = \bar{K}^{(m)}$

$$(3.17)\qquad [\bar{u}_\chi(k)\Psi](x) = \chi[(k_{12}x+k_{22})^{-2}\det k]\Psi\left(\frac{k_{11}x+k_{21}}{k_{12}x+k_{22}}\right)$$

holds whenever $k_{12}x+k_{22}$ is a unit of $\mathfrak{O}/\mathfrak{P}^m$. We shall now prove

LEMMA 3.5. *Let χ be a primitive congruence character* mod $\mathfrak{P}^m$ *and α_1 a primitive congruence character* mod $\mathfrak{P}^{m_1}$. *If $m_1 \neq m$ then $\alpha_1(\overline{AB}) = \alpha_1(\bar{A})$ does not occur in the restriction of $\bar{u}_\chi$ to $\overline{AB}$.*

Proof. We have to consider the restrictions of $\bar{u}_\chi$ to AB. Let us use the explicit form (3.17) of $\bar{u}_\chi$ as multiplier representation on the projective line over $\mathfrak{O}/\mathfrak{P}^m$. If x is a finite point of this line then so is $ax+b$. Hence if Ψ vanishes outside the finite points then $[\bar{u}_\chi(ab)\Psi](x) = \chi(a)\Psi(ax+b)$. Therefore $\bar{u}_\chi(b)\Psi = \Psi$ implies that $\Psi(x) =$ constant for all finite x and for such Ψ clearly $\bar{u}_\chi(ab)\Psi = \chi(a)\Psi$. Next we have to consider the orbits under $\overline{AB}$ consisting of points of infinity. Let $f_l(x) = 0$ unless $\operatorname{ord}(x) = l$ then $f'_l = \bar{u}_\chi(w)f_l$ vanishes outside an orbit under $\overline{AB}$. Now $\bar{u}_\chi(ab)f'_l = \alpha_1(a)f'_l$ is equivalent to $\bar{u}_\chi(w^{-1}abw)f_l = \alpha_1(a)f_l$, i. e.

$$\chi[(a-bx)^{-2}a]f_l\left(\frac{x}{a-bx}\right) = \alpha_1(a)f_l(x)$$

Hence $f_l[\tau^l(1+b\tau^l)] = c\chi(1+b\tau^l)^{-2}$ for all $b \in \mathfrak{O}/\mathfrak{P}^m$ and $\chi(a)^{-1}f_l(a^{-1}\tau^l) = \alpha_1(a)f_l(\tau^l)$ for all units a, in particular for $a^{-1} = 1+b\tau^l$ we get

$$c\chi(1+b\tau^l)^{-2} = c\chi(1+b\tau^l)^{-1}\alpha_1(1+b\tau^l)^{-1}$$

i. e. $\alpha_1(1+b\tau^l) = \chi(1+b\tau^l)$ for all b which is possible only if α_1 is a primitive congruence character mod $\mathfrak{P}^m$.

There only remains one other orbit under $\overline{AB}$, namely the point $1/0$. Its characteristic function Ψ_∞ satisfies $\bar{u}_\chi(ab)\Psi_\infty = \chi(a^{-1})\Psi_\infty$. Thus we have determined all Ψ which satisfy $\bar{u}_\chi(ab)\Psi = c_a\Psi$ and shown that for none of them is c_a equal to $\alpha_1(a)$ which proves Lemma 3.5.

If we combine Lemma 3.5 with equations (3.4), (3.8), (3.9) and (3.10) we obtain

THEOREM 3.3. *Let m_1 be the smallest integer such that $\alpha(1+\mathfrak{P}^{m_1}) = 1$. Let χ be a primitive non-real congruence character* mod $\mathfrak{P}^m$. *If $m \neq m_1$ then u_χ does not occur in the restriction of $M(g,\alpha)$ to K.*

Let us also note

LEMMA 3.6. *The representations*

$$g \to M(g, \tfrac{1}{2}+it, \alpha_1) \text{ and } g \to M(g, \tfrac{1}{2}+it', \alpha_1')$$

of G are equivalent if and only if $(t', \alpha_1') = (t, \alpha_1)$ *or* $(-t, \alpha_1^{-1})$. *The representations* u_χ' *and* u_χ *of* K *are equivalent if and only if* $\chi' = \chi$ *or* χ^{-1}.

4. Computation of the matrix coefficients $M_{jl}(g, \alpha)$. We wish to find a subspace of $L^2(\Omega, dx)$ which transforms according to the representation u_{α_1} of K (see 3.5 above). By (3.11) and Lemma 3.1 above the elements Ψ of this subspace must satisfy $M(k, \alpha)\Psi = \Psi$ for all $k \in K^{(m)}$. Let j vary over a set of representatives for the cosets of $\mathfrak{P}^m$ in $\mathfrak{O}$ and define $\Psi_j(x)$ by

$$\Psi_j(x) = \begin{cases} q^{m/2} & \text{if } x \in \mathfrak{P}^m + j \\ 0 & \text{if } x \notin \mathfrak{P}^m + j. \end{cases} \tag{4.1}$$

Note that these functions are pair-wise orthogonal and

$$\| \Psi_j \|^2 = \int_\Omega | \Psi_j(w) |_\infty^2 \, dx = q^m \int_{\mathfrak{P}^m + j} dx = q^m \int_{\mathfrak{P}^m} dx = 1.$$

Next let w be the linear transformation given by the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, i.e. $wx = -x^{-1}$. If j is as above and a zero divisor mod $\mathfrak{P}^m$, i.e. $j \in \mathfrak{P}$, we put

$$\Psi_{\infty j} = M(w, \alpha)\Psi_j, \text{ i.e. } \Psi_{\infty j}(x) = \begin{cases} \alpha(x)^{-2} q^{m/2} & \text{if } x^{-1} \in \mathfrak{P}^m + j \\ 0 & \text{if } x^{-1} \notin \mathfrak{P}^m + j. \end{cases} \tag{4.2}$$

We note that when α is such that $M(w, \alpha)$ is unitary then $\Psi_{\infty j} \in L^2(\Omega, dx)$ and $\| \Psi_{\infty j} \| = 1$; restricting ourselves to such α we have

$$\langle \Psi_j, \Psi_{j'} \rangle = \delta_{j, j'} \tag{4.3}$$

for all j, j' (including ∞j), where $\langle\, , \rangle$ denotes the inner product in $L^2(\Omega, dx)$. Since j now varies over the union of a set of representatives for the cosets of $\mathfrak{P}^m$ in $\mathfrak{O}$ and the elements $\infty j (= wj$ for $j \in \mathfrak{P}$ (mod $\mathfrak{P}^m$) we obtain an orthonormal set of $q^m + q^{m-1}$ functions Ψ_j for each α for which $M(w, \alpha)$ is unitary.

We show next that $M(k, \alpha)\Psi_j = \Psi_j$ for $k \in K^{(m)}$. Indeed

$$M(k, \alpha)\Psi_j(x) = \alpha(k_{12}x + k_{22})^{-2} \det k \Psi_j \left(\frac{k_{11}x + k_{21}}{k_{12}x + k_{22}} \right).$$

Let $x \in \mathfrak{P}^m + j$ and $j \neq \infty j$, i.e. $j \in \mathfrak{O}$ then $k_{12} + k_{22} \in 1 + \mathfrak{P}^m$, so that $\alpha(k_{12}x + k_{22}) = 1$; also $k_{11}x + k_{21} \in \mathfrak{P}^m + j$, so that if $k \in K^{(m)}$ and $j \neq \infty j$ then $x \in \mathfrak{P}^m + j$ if and only if $(k_{11}x + k_{21})(k_{12}x + k_{22})^{-1} \in \mathfrak{P}^m + j$. This proves $M(k, \alpha)\Psi_j = \Psi_j$ for $k \in K^{(m)}$ and $j \neq \infty, j$. It remains to consider $\Psi_{\infty j}$. We note that for $j \in \mathfrak{P}$ we have

$$M(k,\alpha)\Psi_{\infty j} = M(k,\alpha)M(w,\alpha)\Psi_j = M(w,\alpha)M(w^{-1}kw,\alpha)\Psi_j = M(w,\alpha)\Psi_j$$

since $w^{-1}K^{(m)}w = K^{(m)}$. Hence $M(k,a)\Psi_{\infty j} = \Psi_{\infty j}$ for $k \in K^{(m)}$. Thus the subspace of $L^2(\Omega, dx)$ spanned by the functions $\Psi_j(x)$ (including ∞j) is element-wise invariant under $M(k,\alpha)$ for $k \in K^{(m)}$ and has dimension $q^m + q^{m-1}$. It follows from (3.11) and Lemma 3.1 above that this subspace is invariant under $M(k,\alpha)$ for $k \in K$ and transforms according to u_{α_1} under $M(k,\alpha)$.

For any $g \in G$ we put

$$M_{jl}(g,\alpha) = \langle M(g,\alpha)\Psi_j, \Psi_l\rangle. \tag{4.4}$$

We proceed to compute these $(q^m + q^{m-1})^2$ matrix coefficients for $g \in A$. Let g be given by a matrix $\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$ with $a \in \Omega^\times$. We write (4.5) $a = \tau^n \eta$ where $n = \operatorname{ord}(a)$ and $\eta \in \mathfrak{U}$.

For $j \neq \infty j'$ we have, using (3.15)

$$M(g,\alpha)\Psi_j(x) = \alpha(a)\Psi_j(ax) = \begin{cases} \alpha(a)q^{m/2} & \text{if } ax \in \mathfrak{P}^m + j \\ 0 & \text{if } ax \notin \mathfrak{P}^m + j. \end{cases} \tag{4.6}$$

Let us first consider $\Psi_j(ax)\bar{\Psi}_l(x)$ for $j, l \in \mathfrak{O}$. We put $j = \tau^\mu\epsilon$ and $l = \tau^\nu\epsilon'$ with μ and $\nu < m$ (since j and l are representatives mod $\mathfrak{P}^m$) and $\epsilon, \epsilon' \in \mathfrak{U}$. Now $\Psi_l(x) = 0$ unless $x \in \mathfrak{P}^m + l = \mathfrak{P}^m + \tau^\nu\epsilon'$, i.e. $x = \tau^m\xi + \tau^\nu\epsilon'$ with $\xi \in \mathfrak{O}$. Let us assume $n \geqq 0$ Then $ax = \tau^{m+n}\xi\eta + \tau^{\nu+n}\epsilon'\eta \in \mathfrak{P}^m + \tau^\mu\epsilon$ if and only if $\tau^{\nu+n}\epsilon'\eta - \tau^\mu\epsilon \in \mathfrak{P}^m$. If this is satisfied then $x \in \mathfrak{P}^m + \tau^\nu\epsilon'$ implies $ax \in \mathfrak{P}^m + \tau^\mu\epsilon$ whence

$$M_{jl}(g,\alpha) = \int_{\mathfrak{P}^m + l} \alpha(a) q^m \, dx = \alpha(a).$$

This proves

LEMMA 4.1. *Let* $j = \tau^\mu\epsilon$, $l = \tau^\nu\epsilon'$, $\mu < m$, $\nu < n$, $\epsilon \in \mathfrak{U}$, $\epsilon' \in \mathfrak{U}$ *and* $n = \operatorname{ord} a \geqq 0$. *Then*

$$M_{jl}(g,\alpha) = \begin{cases} \alpha(a) & \text{if } \tau^{\nu+n}\epsilon'\eta - \tau^\mu\epsilon \in \mathfrak{P}^m \\ 0 & \text{if } \tau^{\nu+n}\epsilon'\eta - \tau^\mu\epsilon \notin \mathfrak{P}^m. \end{cases} \tag{4.7}$$

Note that if $j = 0$ this condition becomes $\tau^{\nu+n}\epsilon'\eta \in \mathfrak{P}^m$, i.e. $\nu \geqq m - n$. Hence in particular

$$M_{0l}(g,\alpha) = \begin{cases} \alpha(a) & \text{if } \operatorname{ord}(l) = \nu \geqq m - n \\ 0 & \text{if } \nu < m - n. \end{cases} \tag{4.8}$$

If $l = 0$ then $x = \tau^m\xi$ implies $ax = \tau^{m+n}\xi\eta \in \mathfrak{P}^m$ (since $n \geqq 0$) so that $ax \in \mathfrak{P}^m + j$ only for $j = 0$.

$$M_{j0}(g,\alpha) = \begin{cases} \alpha(a) & \text{if } j = 0 \\ 0 & \text{if } j \neq 0. \end{cases} \tag{4.9}$$

Let us also note that if j is a unit, i.e. $\mu = 0$ then the condition of Lemma 4.1 becomes $\tau^{\nu+n}\xi' \in \mathfrak{P}^m$ which is impossible unless $\nu + n = 0$, i.e. $\nu = n = 0$. Hence in particular

$$(4.10) \qquad M_{jl}(g, \alpha) = 0 \text{ if } j \in \mathfrak{U},\ l \in \mathfrak{P} \text{ and } \mu \geqq 0.$$

It remains to compute $M_{jl}(g, \alpha)$ when j or l are of the form $\infty j'$. Let us consider first the case $j = \infty j'$, i.e. $j = wj'$ with $j' \in \mathfrak{P}$ and $l \in \mathfrak{O}$. We note that

$$\Psi_j(ax) = 0 \text{ unless } (ax)^{-1} \in \mathfrak{P}^m + j', \text{ i.e.}$$
$$1 \leqq -n - \operatorname{ord}(x) \leqq m, \text{ i.e. } -n-1 \geqq \operatorname{ord}(x) \geqq -m-n.$$

But $\Psi_l(x) = 0$ unless $x \in \mathfrak{P}^m + l \subseteq \mathfrak{O}$. Hence if $n = \operatorname{ord} a \geqq 0$ then $\Psi_j(ax)\Psi_l(x) = 0$ for all x. This proves

$$(4.11) \qquad M_{jl}(g, \alpha) = 0 \text{ if } j = wj',\ l \in \mathfrak{O} \text{ and } \operatorname{ord}(a) \geqq 0.$$

Now we compute $M_{jl}(g, \alpha)$ for $j \in \mathfrak{O}$ and $l = wj'$, $j' \in \mathfrak{O}$. By (4.1) and (4.2) we have

$$(4.12) \quad \alpha(a)\Psi_j(ax)\Psi_l(x) = \begin{cases} 0 \text{ if } ax\epsilon \notin \mathfrak{P}^m + j \text{ or } x^{-1} \notin \mathfrak{P}^m + j' \\ q^m \alpha(a)\alpha(x)^{-2} \text{ if } ax \in \mathfrak{P}^m + j \text{ and } x^{-1} \in \mathfrak{P}^m + j'. \end{cases}$$

Thus we have to consider $x \in [a^{-1}(\mathfrak{P}^m + j)] \cap [\mathfrak{P}^m + j']^{-1}$. Let us write $j = \tau^\mu \epsilon$, $j' = \tau^\nu \epsilon'$ with $\epsilon, \epsilon' \in \mathfrak{U}$ and $0 \leqq \mu, \nu < m$. Then $x \in a^{-1}(\mathfrak{P}^m + j)$ means $x = \tau^{-n}\eta^{-1}(\tau^m \xi + \tau^\mu \epsilon) = \tau^{\mu-n}\eta^{-1}(\epsilon + \tau^{m-\mu}\xi)$ with $\xi \in \mathfrak{O}$. On the other hand $x \in (\mathfrak{P}^m + j')^{-1}$ means $x^{-1} = \tau^m \xi' + \tau^\nu \epsilon' = \tau^\nu(\epsilon' + \tau^{m-\nu}\xi')$ with $\xi' \in \mathfrak{O}$, i.e. $x = \tau^{-\nu}(\epsilon' + \tau^{m-\nu}\xi')^{-1} = \tau^{-\nu}\epsilon'^{-1}(1 + \tau^{m-\nu}\xi')^{-1}$ with $\zeta' = \epsilon'^{-1}\xi'$. So we must have $\mu = n - \nu$, $\epsilon' = \eta\epsilon^{-1}$ and with $\zeta = \epsilon^{-1}\xi$

$$(1 + \tau^{m-\mu}\zeta) = (1 + \tau^{m-\nu}\zeta')^{-1} \text{ i.e. } \tau^{m-\mu}\zeta + \tau^{m-\nu}\zeta' + \tau^{2m-\mu-\nu}\zeta\zeta' = 0,$$

hence $\zeta' = -\tau^{-\mu+\nu}\zeta(1 + \zeta^{m-\mu})^{-1}$. Since $\zeta \in \mathfrak{O}$ and $\zeta' \in \mathfrak{O}$ and $\mu < m$ we must have $\operatorname{ord} \xi \geqq \max\{0, \mu - \nu\}$. In particular we have proved that

$$(4.13) \qquad \begin{aligned} &M_{jl}(g, \alpha) = 0 \text{ if } l = wj',\ j = \epsilon\tau^\mu,\ j' = \epsilon'\tau^\nu \text{ with } \epsilon', \epsilon \in \mathfrak{U} \text{ and} \\ &0 \leqq \mu, \nu < m \text{ unless } \mu = n - \nu,\ \epsilon' = \eta\epsilon^{-1}. \end{aligned}$$

So let us assume now $\mu = n - \nu$, $\epsilon' = \eta\epsilon^{-1}$. Then

$$x = \tau^{-\nu}\eta^{-1}\epsilon(1 + \tau^{m-\mu}\xi) \text{ with } \begin{cases} \xi \in \mathfrak{O} & \text{if } \mu \leqq \nu \\ \xi \in \mathfrak{P}^{\mu-\nu} & \text{if } \mu > \nu. \end{cases}$$

Thus

$$x = \begin{cases} \tau^{-\nu}\eta^{-1}\epsilon(1 + \tau^{m-\nu}\xi) & \text{if } \mu \leqq \nu \\ \tau^{-\nu}\eta^{-1}\epsilon(1 + \tau^{m-\mu}\xi) & \text{if } \mu > \nu \end{cases}$$

with ξ varying over $\mathfrak{O}$. Let us consider the case $\mu \leqq \nu$. Then

$$\alpha(x) = \alpha(\tau^{\mu} a^{-1} \epsilon)\alpha(1 + \tau^{m-\mu}\xi) \text{ with } \xi \in \mathfrak{D}.$$

If $\mu = 0$ then $\alpha(1 + \tau^m \xi) = 1$ since α_1 is a character mod $\mathfrak{P}^m$, so that in this case

$$\int \alpha(a)\psi_j(ax)\bar{\psi}_l(x)\,dx = \alpha(a) q^m \bar{\alpha}(a^{-1}\epsilon)^{-2} \int d[a^{-1}\epsilon(1 + \tau^m \xi)]$$

$$= q^m \alpha(a)\bar{\alpha}(a^2)\alpha(\epsilon^2)\,|\,a^{-1}\,|\int_{\mathfrak{P}^m} d\xi = \alpha(a^{-1})\alpha(\epsilon^2) q^{-n} \text{ for } Rs = \tfrac{1}{2}.$$

If $\mu > 0$ we get

$$\alpha(a) q^m \bar{\alpha}(a^{-1}\epsilon\tau^{\mu})^{-2} \int_{\mathfrak{D}} \bar{\alpha}(1 + \tau^{m-\mu}\xi)^{-2} d[a^{-1}\epsilon\tau^{\mu}(1 + \tau^{m-\mu}\xi)]$$

$$= \bar{\alpha}(\tau^{-2\mu})\alpha(a^{-1})\alpha(\epsilon^2) q^{m-n-\mu} \int_{\mathfrak{P}^{m-\mu}} \alpha(1 + \xi^2)\,d\xi \text{ for } Rs = \tfrac{1}{2},$$

so that for $0 < \mu \leqq \nu$ we get

$$(4.15) \qquad M_{jl}(g, \alpha) = \begin{cases} 0 & \text{if } \alpha(1 + \mathfrak{P}^{m-\mu})^2 \neq 1 \\ \alpha(a^{-1})\bar{\alpha}(\tau^{-2\mu})\alpha(\epsilon^2) q^{-n} & \text{if } \alpha(1 + \mathfrak{P}^{m-\mu})^2 = 1. \end{cases}$$

Note that α_1 is a primitive non-real character mod $\mathfrak{P}^m$, hence so is α_1^2 unless q is a power of 2. Hence

$$(4.16) \qquad M_{jl}(g, \alpha) = 0 \text{ if } 0 < \mu \leqq \nu \text{ and } (2, q) = 1.$$

The case $\mu > \nu$ is the same as the case $\mu \leqq \nu$ with μ replaced by ν. Therefore it remains only to consider $j = 0$ or $j' = 0$.

If both $j = 0$ and $j' = 0$ then we have

$$M_{0\infty}(g, \alpha) = \int \alpha(a)\psi_0(ax)\bar{\psi}_{\infty}(x)\,dx = \alpha(a) q^m \int_{1-n \leqq \operatorname{ord} x \leqq -m} \bar{\alpha}(x^{-2})\,dx$$

$$= \alpha(a) q^{m-\nu} \sum_{\nu=1-n}^{-m} \bar{\alpha}(\tau^{-2\nu}) \int_{\mathfrak{U}} \alpha(\eta)^2\,d\eta = 0$$

since $\alpha_1^2 \neq 1$.

If $j = \epsilon\tau^{\mu}$ $(0 \leqq \mu < m)$ and $j' = 0$, then ws must have $ax \in \mathfrak{P}^m + j$ and $x^{-1} \in \mathfrak{P}^m$; but $x = a^{-1}(\tau^m \xi + \epsilon\tau^{\mu}) = a^{-1}\tau^{\mu}(\epsilon + \tau^{m-\mu}\xi)$ with $\xi \in \mathfrak{D}$ satisfies $x^{-1} \in \mathfrak{P}^m$ if and only if $\mu \leqq n - m$. Assuming this we get

$$\alpha(a) q^m \int \bar{\alpha}(x^{-2})\,dx = \alpha(a) q^m \bar{\alpha}(a^{-1}\tau^{\mu})^{-2} \int \bar{\alpha}(\epsilon + \tau^{m-\mu}\xi)^{-2} d[a^{-1}\tau^{\mu}(\epsilon + \tau^{m-\mu}\xi)]$$

$$= \alpha(a)\bar{\alpha}(a^2)\bar{\alpha}(\tau^{-2\mu}) q^n \alpha(\epsilon^2) \int_{\mathfrak{D}} \alpha(1 + \tau^{m-\mu}\xi)^2 d\xi$$

$$= \bar{\alpha}(\tau^{-2\mu})\alpha(a^{-1}) q^{-n}\alpha(\epsilon^2) \int_{\mathfrak{D}} \alpha(1 + \tau^{m-\mu}\xi)^2 d\xi.$$

Hence if $j = \epsilon\tau^{\mu}$ $(0 \leqq \mu < m)$ and $j' = 0$ we have

$$(4.17)\quad M_{j\infty}(g,\alpha)=\begin{cases} 0 & \text{if } \mu > n-m \text{ or if } \alpha(1+\mathfrak{P}^{m-\mu})^2\neq 1 \\ \alpha(a^{-1})\bar{\alpha}(\tau^{-2\mu})\alpha(\epsilon^2)q^{-n} & \text{if } \mu\leqq n-m \text{ and } \alpha(1+\mathfrak{P}^{m-\mu})^2=1. \end{cases}$$

Hence if $n=\operatorname{ord} a<m$ then $M_{j\infty}(g,\alpha)=0$. Also if q is not a power of 2, then $\alpha(1+\mathfrak{P}^{m-\mu})^2\neq 1$ since α_1 is a primitive non-real character mod $\mathfrak{P}^m$. This proves

LEMMA 4.2. *Let $j=\epsilon\tau^\mu$, $l=wj'$, $j'=\epsilon'\tau^\nu$ with $\epsilon,\epsilon'\in\mathfrak{U}$, $0\leqq\mu<m$ and $1\leqq\nu<m$. If $(2,q)=1$ then*

$$M_{jl}(g,\alpha)=\begin{cases}\alpha(a^{-1})\alpha(\epsilon^2)q^{-n} & \textit{if } n=\nu,\ \epsilon'=\eta\epsilon^{-1} \textit{ and } \mu=0.\\ 0 & \textit{otherwise.}\end{cases}$$

If q is a power of 2 the result remains the same when $\mu=0$.

If $0<\mu\leqq\nu$ then $M_{jl}(g,\alpha)=\alpha(a^{-1})\bar{\alpha}(\tau^{-2\mu})\alpha(\epsilon^2)q^{-n}$ or 0 according as $\alpha(1+\mathfrak{P}^{m-\mu})^2=1$ or not. If $0<\mu<\nu$ then $M_{jl}(g,\alpha)=\alpha(a^{-1})\bar{\alpha}(\tau^{-2\mu}\alpha(\epsilon^2)q^{-n}$ or 0 according as $\alpha(1+\mathfrak{P}^{m\nu})^2=1$ or not.

If $j=j'=0$ then $M_{0\infty}(g,\alpha)=0$ since $\alpha_1{}^2\neq 1$. If $j=\epsilon\tau^\mu$ and $j'=0$, then $M_{j\infty}(g,\alpha)$ is given by (4.17).

From this we conclude at once

LEMMA 4.3. *Let j and l be as in Lemma 4.2 above. Then*

$$M_{lj}(g,\alpha)=0 \textit{ if } \mathrm{n}=\operatorname{ord} a>0.$$

Proof. We may assume α such that $M(g,\alpha)$ is unitary. Then $M_{lj}(g,\alpha)=\overline{M_{jl}(g^{-1},\alpha)}$ so that we are reduced to the case $n<0$ of above Lemma 4.2.

It remains to consider the case l, l' with $l=wj$, $l'=wj'$, $j=\epsilon\tau^\mu$, $j'=\epsilon'\tau^\nu$. Then

$$\langle M(g,\alpha)\psi_l,\psi_{l'}\rangle=\langle M(g,\alpha)M(w,\alpha)\psi_j,\ M(w,\alpha)\psi_{j'}\rangle=\langle M(g^{-1},\alpha)\psi_j,\psi_{j'}\rangle$$

so that we are reduced to the case $M_{jj'}$ with $n<0$, and $j,j'\in\mathfrak{O}$. As before, $\psi_{j'}(x)=0$ unless $x\in\mathfrak{P}^m+\epsilon'\tau^\nu$, i.e. $x=\tau^m\xi+\tau^\nu\epsilon'$ with $\xi\in\mathfrak{O}$. But $\psi_j(ax)=0$ unless $ax\in\mathfrak{P}^m+\epsilon\tau^\mu$ i.e. $x=\eta^{-1}\tau^{m-n}\xi^1+z^{-1}\epsilon\tau^{\mu-n}$ with $\xi'\in\mathfrak{O}$. Since $n<0$ now, $x\in\mathfrak{P}^m+\epsilon'\tau^\nu$ implies $x\in\mathfrak{P}^{m-n}+\epsilon'\tau^\nu$. Therefore $\psi_j(ax)\bar{\psi}_{j'}(x)=0$ unless $\epsilon\eta^{-1}\tau^{\mu-n}-\epsilon'\tau^\nu\in\mathfrak{P}^m$. Assuming this we obtain for $n<0$

$$M_{jj'}(g,\alpha)=\int\alpha(a)q^m\,dx=\alpha(a)q^m\int d(\eta^{-1}\tau^{m-n}\xi+\eta^{-1}\epsilon\tau^{\mu-n})=\alpha(a)q^n$$

This proves

LEMMA 4.4. *Let $l=wj$, $l'=wj'$, $j,j'\in\mathfrak{P}$, we write $j=\epsilon\tau^\mu$, $j'=\epsilon'\tau^\nu$ with μ and $\nu\geqq 1$. Then for $n=\operatorname{ord} a>0$ we have*

$$M_{ll'}(g,\alpha) = \begin{cases} \alpha(a)^{-1}q^{-n} \text{ if } \epsilon\eta^{-1}\tau^{\mu-n} - \epsilon'\tau^{\nu} \in \mathfrak{P}^m \\ 0 \text{ if not.} \end{cases}$$

THEOREM 4.1. *$M_{jl}(g,\alpha)$ is given by the following matrix*[5]

$$\begin{array}{cc}
\begin{array}{r}
\mathfrak{P}\left\{\begin{array}{c} \\ \\ \end{array}\right. \\
\mathfrak{U}\left\{\begin{array}{c} \\ \\ \end{array}\right. \\
\infty\left\{\begin{array}{c} \\ \\ \end{array}\right.
\end{array}
&
\left[\begin{array}{lll}
\cdots\alpha(a) & \text{or, } 0 & \quad\vdots \\
 & & \alpha(a)^{-1}\alpha(\epsilon)^2q^{-n} \\
0\cdots0\cdots & 0\cdots0 & \cdot\text{ or } 0 \\
0\cdots0\cdots & 0\cdots0 & \quad\vdots \\
0\cdots0\cdots & 0\cdots0 & \quad\vdots \\
0\cdots0\cdots & 0\cdots0 & \alpha(a)^{-1}q^{-n} \text{ or } 0
\end{array}\right]
\\
 & \begin{array}{ccc} \underbrace{\qquad\quad}_{\mathfrak{P}} & \underbrace{\qquad\quad}_{\mathfrak{U}} & \underbrace{\qquad\qquad\quad}_{\infty} \end{array}
\end{array}$$

5. **The mapping $f \to F(\alpha)$.** We now wish to study the integral

$$\int_G f(g)M(g,\alpha)^*\,dg \tag{5.1}$$

for $f \in S_e^0$. We introduce in the representation space of $M(g,\alpha)$ a complete orthonormal set which is adapted to the decomposition under $M(k,\alpha)$, $k \in K$, i. e. we decompose the Hilbert space under $M(k,\alpha)$ into a direct sum of pairwise orthogonal invariant irreducible subspaces and introduce an orthonormal base in each of these subspaces; their union is the complete orthonormal base which we shall use.

The representation $u = u_{\alpha_1}$ of K introduced in § 3 above occurs exactly once in the decomposition under $M(k,\alpha)$, k $\in K$, as we have seen in § 3 above (see Theorem 3.1). Let us denote the corresponding subspace by $\mathfrak{v}$. It follows from the orthogonality relations on K that all the matrix coefficients of $\mathfrak{F}(\alpha) = \int f(g)M(g,\alpha)^*\,dg$ vanish except possibly

$$\langle\mathfrak{F}(\alpha)\Psi_j, \Psi_l\rangle \text{ for } \Psi_j \text{ and } \Psi_l \in \mathfrak{v}. \tag{5.2}$$

Now

$$\langle\mathfrak{F}(\alpha)\Psi_j, \Psi_l\rangle = \int_G f(g)\,\bar{M}_{lj}(g,\alpha)\,dg$$

where

$$M_{jl}(g,\alpha) = \langle M(g,\alpha)\Psi_j, \Psi_l\rangle. \tag{5.3}$$

If $e(k) = ru_{kk}(k) = r\langle M(k,\alpha)\Psi_k, \Psi_k\rangle$ in accordance with (2.1), then orthogonality relations on K imply moreover that

---

[5] If q is a power of 2 then $\alpha(a^{-1})\alpha(\tau^{-2\mu})\alpha(\epsilon^2)q^{-n}$ can occur instead of $\alpha(a^{-1})\alpha(\epsilon^2)q^{-n}$ in the last column.

$$(5.4) \qquad \langle \mathfrak{F}(\alpha)\Psi_j, \Psi_l\rangle = 0 \text{ if } (j,l) \neq (i,i).$$

Hence there remains but one non-zero matrix-coefficient namely

$$(5.5) \qquad \langle \mathfrak{F}(\alpha)\Psi_i, \Psi_i\rangle = \int_G f(g)\bar{M}_{ii}(g,\alpha)\,dg.$$

It follows that the mapping

$$(5.6) \qquad f \rightarrow \langle \mathfrak{F}(\alpha)\Psi_i, \Psi_i\rangle$$

is for each fixed α a homomorphism of the algebra $S_e{}^0$ to the complex numbers.

We wish to evaluate the integral on the right side of (5.5). For this purpose let us make the following observations. Let $\mathfrak{h}(g) \in S_u$ and let $h_{jl}(g)$ be its matrix coefficients with respect to the above basis in $\mathfrak{b}$; we assume that $h_{jl}(g)$ is integrable on compact sets. By Lemma 2.1 there exists an element $\mathfrak{f}(g)$ of $S_u{}^0$ whose (i,i)-th matrix coefficient is $1/r$ times the above $f(g) \in S_u{}^0$:

$$(5.7) \qquad \mathfrak{f}(g) = [f_{jl}(g)] \text{ with } rf_{ii}(g) = f(g).$$

Hence

$$\int_G f(g)\bar{h}_{ii}(g)\,dg = \int_K \int_K \sum_{n=0}^{\infty} rf_{ii}(ky^n\mathfrak{k})\bar{h}_{ii}(ky^n\mathfrak{k})\,dk\,d\mathfrak{k}\,\mathfrak{M}(Ky^nK)$$
$$= r\int_K \int_K \sum_n \sum_{l,m} u_{il}(k)f_{lm}(y^n)u_{mi}(\mathfrak{k}) \sum_{\lambda,\mu} \bar{u}_{i\lambda}(k)\bar{h}_{\lambda\mu}(y^n)\bar{u}_{\mu i}(\mathfrak{k})\,dk\,d\mathfrak{k}\,\mathfrak{M}(Ky^nK).$$

using the definition of S_u, i.e. (1.1) and the normalization of the measure on K: $\mathfrak{M}(K) = 1$. Now let us apply the orthogonality relations on K and we obtain

$$(5.8) \qquad \int_G f(g)\bar{h}_{ii}(g)\,dg = r^{-1}\sum_{n=0}^{\infty} \sum_{j,l} f_{j,l}(y^n)\bar{h}_{jl}(y^n)\mathfrak{M}(Ky^nK)$$

this proves

LEMMA 5.1. *Let* $f(g) \in S_e{}^0$ *and* $\mathfrak{h}(g) \in S_u$ *with* $\mathfrak{h}(g) = [h_{lm}(g)]$ *where* $h_{lm}(g)$ *is integrable on compact sets. Then* (5.8) *holds. Hence in particular* $\int f(g)\bar{h}_{ii}(g)\,dg$ *is independent of* i.

We apply this with $h_{jl}(g) = M_{jl}(g,\alpha)$. The conditions of Lemma 5.1 are clearly satisfied, in particular if $M(g,\alpha)$ is unitary then $M_{jl}(g,\alpha)$ is even a bounded continuous function on G. As we now know that (5.5) is independent of i, we may write for any $f(g) \in S_e{}^0$

$$(5.9) \qquad F(\alpha) = \langle \mathfrak{F}(\alpha)\Psi_i, \Psi_i\rangle = \int f(g)\bar{M}_{ii}(g,\alpha)\,dg.$$

So far we have only used in this section that the irreducible representation u of K occurs exactly once in $M(k,\alpha)$. Now let us use the explicit representation $u = u_{\alpha_1}$ discussed in § 3 (see in particular (3.3) and (3.5) above) and let us insert the explicit expression for $M_{jl}(g, \alpha)$ computed in § 4 (see Theorem 4.1). We set $g = y^n$, i.e. $a = \tau^n$ and obtain for any $f(g) \in S_e^0$, using Lemma 5.1.

$$\begin{aligned} F(\alpha) &= r^{-1}\{ \sum_{j=1}^{r} f_{jj}(1) + \sum_{n=1}^{\infty} \sum_{j,l=1}^{r} f_{jl}(y^n)\, \bar{M}_{lj}(y^n, \alpha) \mathfrak{M}(Ky^nK) \} \\ &= r^{-1}\{ \sum_{j=1}^{r} f_{jj}(1) + \sum_{n=1}^{\infty} [\bar{\alpha}(\tau^n) \sum_{(j,l)\in I} f_{jl}(y^n) \\ &\quad + \bar{\alpha}(\tau^n) q^{-n} \sum_{(j,l)\in II} f_{jl}(y^n) \alpha(\epsilon_j)^2 \\ &\quad + \bar{\alpha}(\tau^{-n}) q^{-n} \sum_{(j,l)\in III} f_{jl}(y^n)] q^{n-1}(q+1) \} \end{aligned}$$

using $\mathfrak{M}(Ky^nK) = q^{n-1}(q+1)$ for $n \geqq 1$, $\mathfrak{M}(K) = 1$ and the above expressions for $M_{jl}(y^n, \alpha)$ (see § 4 above). Let us note that $r = q^m + q^{m-1}$ and α is such that $M(g, \alpha)$ is unitary. Hence by Lemmas 3.2 and 3.3 α must be of the form $\alpha(\tau^n\eta) = q^{n(it-\frac{1}{2})}\alpha_1(\eta)$; moreover α_1 is a fixed character of $\mathfrak{U}$. Hence, writing

$$F(\alpha) = F(\tfrac{1}{2} + it) \text{ when } \alpha(\tau^n\eta) = q^{n(it-\frac{1}{2})}\alpha_1(\eta), \tag{5.10}$$

we obtain

$$\begin{aligned} &(q^m + q^{m-1}) F(\tfrac{1}{2} + it) - \sum_j f_{jj}(1) \\ &= \sum_{n=1}^{\infty} [q^{-n(it+\frac{1}{2})} \sum_{(j,l)\in I} f_{jl}(y^n) + q^{n(it-\frac{1}{2})} \sum_{(j,l)\in II} f_{jl}(y^n) \alpha(\epsilon_j)^2 \\ &\qquad + q^{n(it-\frac{1}{2})} \sum_{(j,l)\in III} f_{jl}(y^n)] q^{n-1}(q+1). \end{aligned} \tag{5.11}$$

Here $(j, l) \in I$ means that j, l varies over those indices for which $M_{jl}(y^n, \alpha) = \alpha(y^n)$ (see Lemma 4.1), similarly $(j, l) \in II$ if and only if $M_{jl}(y^n, \alpha) = \alpha(a)^{-1} q^{-n} \alpha(\epsilon)^2$ as given by Lemma 4.2 and $(j, l) \in III$ if and only if $M_{jl}(y^n, \alpha) = \alpha(y^{-n}) q^{-n}$ as given by Lemma 4.3. Hence writing

$$\begin{aligned} f_I(y^n) &= \sum_{(j,l)\in I} f_{jl}(y^n), \quad f_{II}(y^n) = \sum_{(j,l)\in II} f_{jl}(y^n) \alpha(\epsilon_j)^2 \\ f_{III}(y^n) &= \sum_{(j,l)\in III} f_{jl}(y^n) \end{aligned} \tag{5.12}$$

we obtain

$$\begin{aligned} &(q^m + q^{m-1}) F(\tfrac{1}{2} + it) - f(1) \\ &= \sum_{n=1}^{\infty} [q^{-n(\frac{1}{2}+it)} f_I(y^n) + q^{-n(\frac{1}{2}-it)} f_{III}(y^n) + q^{n(it-\frac{1}{2})} f_{II}(y^n)] q^{n-1}(q+1) \\ &= \sum_{n=1}^{\infty} q^{n-1}(q+1) q^{-nit} q^{-n/2} f_I(y^n) + q^{nit} q^{n/2} [f_{II}(y^n) + f_{III}(y^n)] q^{n-1}(q + \end{aligned} \tag{5.13}$$

noting that $f_{jj}(1) = r^{-1}f(1)$ so that $\sum_{j=1}^{r} f_{jj}(1) = f(1)$ for $f \in S_e^2$ with $f_{ji}(g)$ defined by (5.7). If q is a power of 2 the set II has to be split up in accordance with Lemma 4.2 and expressions (5.13) and (5.13) have to be changed accordingly.

6. The Plancherel formula for the subspace S_e^2. For fixed α_1 and varying t we have thus found an explicit expression for the map $f \to F(\frac{1}{2}+it)$ for any $f \in S_e^0$, as given by (5.13). We shall now introduce a subspace of S_e^2 on which this map is naturally one-one.

Definition 6.1. Let $\mathfrak{S}_e^2$ be the orthogonal complement in S_e^2 of the kernel of the map $f \to F(\frac{1}{2}+it)$.

We extend the map $f \to F(\frac{1}{2}+it)$ in the usual manner to S_e^2 and then obtain

THEOREM 6.1. *Let $\bar{u} = \bar{u}_\chi$ be a representation in the principal series* mod $\mathfrak{P}^m$ *corresponding to a primitive non real congruence character* χ mod q^m. *Let u be the corresponding representation of K (see* § 3 *especially* (3.3) *and* (3.5)) *and $e(k) = ru_{ii}(k)$ for any fixed $i = 1, 2, \cdots, r$ (see* (2.1) *above). Then the map $f \to F(\frac{1}{2}+it)$ is a unitary operator of the Hilbert space $\mathfrak{S}_e^2$ onto $L^2(0 \leqq t \leqq 2\pi/\log q;\ (q^m+q^{m-1})\,dt \log q/2\pi)$. The inverse is given by.*

$$f(g) = \int_0^{2\pi/\log q} F(\tfrac{1}{2}+it) M_u(g, \tfrac{1}{2}+it)\, dt \log q.(q^m+q^{m-1})/2\pi. \tag{6.1}$$

For any function $F(\frac{1}{2}+it) \in L^2$ $(0 \leqq t \leqq 2\pi/\log q;\ dt)$ the integral (6.1) *exists and defines a function $f(g) \in \mathfrak{S}_e^2$.*

Proof. As throughout α_1 denotes the restriction of α to $A \cap K \cong \mathfrak{U}$. Let m_1 be the smallest integer such that $\alpha_1(1+\mathfrak{P}^{m_1}) = 1$. If χ is a primitive congruence character mod $\mathfrak{P}^m$ and if $m \neq m_1$ then by Theorem 3.3 u_χ does not occur in the restriction $M(K, \alpha)$ of $M(g, \alpha)$ to K. This implies that $F(\alpha) = 0$ for any such α. There remain only those α for which $m_1 = m$. By Theorem 3.1 $\bar{u}_{\alpha_1}$ and $\bar{u}_\chi$ must be equivalent representations, otherwise $u_\chi \not\subseteq M(K, \alpha)$. By Lemma 3.6 $\bar{u}_{\alpha_1}$ and $\bar{u}_\chi$ are equivalent representations if and only if $\alpha_1 = \chi$ or $\alpha_1 = \chi^{-1}$. It follows that for $f \in S_e^2$

$$F(\alpha) = 0 \text{ unless } m_1 = m \text{ and } \alpha_1 = \chi \text{ or } \chi^{-1}. \tag{6.2}$$

Hence we let $\alpha_1 = \chi$ be fixed and let t vary over the real interval $0 \leqq t \leqq 2\pi/\log q$. Now we use the explicit form (5.13) of the map $f \to F(\frac{1}{2}+it)$ which shows that $F(\frac{1}{2}+it)$ is a finite Fourier series whenever

$f \in S_e^0$ and that this map is a bounded operator of S_e^2 onto $L^2(0 \leqq t \leqq 2\pi/\log q; dt)$. The explicit form (5.13) also shows that this map is unitary on S_e^2 and in fact gives an explicit description of $\mathfrak{S}_e^2$.

Indeed if $F(\alpha) = 0$ for all α then by (5.13) we must have

$$f_I(y^n) = f_{II}(y^n) = f_{III}(y^n) = 0 \text{ for all } n \geqq 1 \tag{6.3}$$

and conversely. By definition of $\mathfrak{S}_e$ any $h \in \mathfrak{S}_e$ must be orthogonal in $L^2(G)$ to all $f \in S_e^2$ which satisfy (6.3). Using Lemma 5.1 we obtain the following description of $\mathfrak{S}_e^2$.

LEMMA 6.1. *Let $f \in S_e^2$ and let us form the matrix valued function $f_{ij}(g)$ as in Lemma 2.1. Let $(2, q) = 1$. Then $f \in \mathfrak{S}_e^2$ if and only if*

$$f_{ij}(y^n) = f_{lm}(y^n) \text{ for all } n \geqq 1 \tag{6.4}$$

whenever (i, j) and (l, m) both belong to I, or both belong to III; if both (i, j) and (l, m) belong to II then

$$\alpha(\epsilon_i)^{-2} f_{ij}(y^n) = \alpha(\epsilon_l)^{-2} f_{lm}(y^n) \text{ for } n \geqq 1. \tag{6.4a}$$

If $(i, j) \in II$ and $(l, m) \in III$

$$f_{ij}(y^n) = \alpha(\epsilon_i)^2 f_{lm}(y^n) \text{ for all } n \geqq 1. \tag{6.5}$$

The set of indices I, II, III are defined in § 5 after (5.11). Moreover if i, j does not belong to I, II or III then

Moreover if i, j does not belong to I, II or III then

$$f_{ij}(y^n) = 0 \text{ for all } n \geqq 1. \tag{6.6}$$

Now let us integrate (5.13) with respect to t and we obtain

$$f(1) = (q^m + q^{m-1}) \int_0^{2\pi/\log q} F(\tfrac{1}{2} + it)\, dt \log q/2\pi. \tag{6.7}$$

This implies the inversion formula (6.1) and hence the fact that the map $f \to F(\frac{1}{2} + it)$ is a unitary map onto

$$L^2(0 \leqq t \leqq 2\pi/\log q;\ dt(q^m + q^{m-1}) \log q/2\pi).$$

This concludes the proof of Theorem 6.1.

7. The subalgebra $\mathfrak{S}_e^0$. Let $\mathfrak{S}_e^0 = \mathfrak{S}_e^2 \cap S_e^0$. One sees easily that $\mathfrak{S}_e^0$ is a subalgebra of S_e^0 (with convolution as product). In fact we have

THEOREM 7.1. *Let u satisfy hypotheses of Theorem 6.2. Then the map*

*$f \to F(\frac{1}{2}+it)$ is an isomorphism of the subalgebra $\mathfrak{S}_e{}^0$ onto the algebra of all Fourier polynomials $\sum c_n q^{int}$. Hence every homomorphism of $\mathfrak{S}_e{}^0$ to the field of complex numbers is of the form $f \to F(s)$ for some complex number s. This is a *-homomorphism if and only if $Rs = \frac{1}{2}$.*

Proof. By Theorem 6.1 the map $f \to F(\frac{1}{2}+it)$ is one-one on $\mathfrak{S}_e{}^2$, hence on $\mathfrak{S}_e{}^0$. By the general theory of group representation it is homomorphism, hence an isomorphism. To prove that the image consists of all Fourier polynomials one only has to use the explicit form (5.13) together with the explicit computation of the matrix-coefficients $M_{jl}(g,\alpha)$ (see § 4) in order to write down for each power q^{nit} an explicit inverse image $f \in \mathfrak{S}_e{}^0$.

Indeed by Lemma 6.1 we know that $\mathfrak{S}_e{}^0$ consists exactly of those $f \in S_e{}^0$ which satisfy (6.4) to (6.6). Let such an f also satisfy $f_I(y^n) = 0$ for all $n \geqq 0$ and $f_{III}(y^n) = 0$ for all but one value of $n \geqq 1$ then the corresponding $F(s)$ is a constant multiple of $q^{n(s-\frac{1}{2})}$ for this $n \geqq 1$ and such elements of $S_e{}^0$ exist by Lemma 6.1. Similarly for $n \leqq 1$.

Let us also note that if $F(s)$ corresponds to $f \in \mathfrak{S}_e{}^0$ then by (5.13) there corresponds to f^* the Fourier polynomial $\bar{F}(1-\bar{s})$. Hence the *-homomorphisms of the algebra $\mathfrak{S}_e{}^0$ corresponds to the points s with $Rs = \frac{1}{2}$.

Let us denote by $\mathfrak{S}_e(\nu)$ the elements f of $\mathfrak{S}_e$ which vanish outside the double coset $Ky^\nu K$, i.e. we write

$$\mathfrak{S}_e(\nu) = \mathfrak{S}_e \cap S_e(\nu). \tag{7.1}$$

We now have easily

THEOREM 7.2. *$\mathfrak{S}_e(\nu)$ is a 2-dimensional vector space over the complex numbers for $\nu \geqq 1$, if q is not a power of 2.*

This is now an immediate consequence of Lemma 6.1.

Definition 7.1. Let us denote by $\mathfrak{S}_u{}^0$ the image of $\mathfrak{S}_e{}^0$ under the isomorphism (2.4) of Lemma 2.1. Let $\mathfrak{S}_u(\nu)$ be the space of elements of $\mathfrak{S}_u{}^0$ which vanish outside the double coset $Ky^\nu K$.

We note that $\mathfrak{S}_u(\nu)$ is mapped onto $\mathfrak{S}_e(\nu)$ under the isomorphism (2.4) as follows easily from the above.

Let us also write

$$\mathfrak{S}_u = \bigcup_{\nu=0}^{\infty} \mathfrak{S}_u(\nu). \tag{7.2}$$

8. $\mathfrak{S}_e{}^0 \neq S_e{}^0$. We shall now study the space S_u in some detail, when u is the principal series mod $\mathfrak{P}^m$. Let ν be a fixed integer > 0 and consider the

double coset $Ky^\nu K$. We shall have to know how many linearly independent elements of S_u vanish outside $Ky^\nu K$. By § 2 above this is equivalent with the determination of all matrices X which satisfy

$$(8.1) \qquad u(k)X = Xu(\kappa) \text{ whenever } ky^\nu = y^\nu\kappa \text{ for } k \text{ and } \kappa \in K.$$

Any such k must belong to K_ν where K_ν is the image under the natural homomorphism $GL(2,\Omega) \rightarrow PGL(2,\Omega)$ of the group of matrices (k_{ij}) with $k_{ij} \in \mathfrak{O}$, $k_{12} \in \mathfrak{P}^\nu$ and $k_{11}k_{22} - k_{12}k_{21} \in \mathfrak{U}$. Then k and κ vary in (8.1) over all pairs where $k \in K_\nu$ and $\kappa = y^{-\nu}ky^\nu$. Let

(8.2) $u_\nu(k) = u(y^{-\nu}ky^\nu)$ for $k \in K_\nu$
then u_ν is a representation of K_ν and (8.1) is equivalent to

(8.3) $u(k)X = Xu_\nu(k)$ for all $k \in K_\nu$.
Thus X varies over all intertwining operators between the representations u and u_ν of K_ν.

Now we use the explicit form of u as principal series corresponding to a character $\chi \bmod \mathfrak{P}^m$, namely

$$(8.4) \qquad [u(k)\psi](x) = \chi[(k_{12}x + k_{22})^{-2} \det k]\psi\left(\frac{k_{11}x + k_{21}}{k_{12}x + k_{22}}\right)$$

where x varies over the points of the projective line over $\mathfrak{O}/\mathfrak{P}^m$ and the right side has meaning whenever $(k_{12}x + k_{22})^{-2} \det k \notin \mathfrak{P}$.

Let $k \in K_\nu$ and $\kappa = y^{-\nu}ky^\nu$, then κ is given by a matrix (κ_{ij}) with $\kappa_{21} \in \mathfrak{P}^\nu$. For any positive integer $l < m$ let the function $f_l(x)$ be defined by

$$(8.5) \qquad f_l(x) = \begin{cases} \chi(\epsilon)^{-2} & \text{if } x = \epsilon\tau^l \\ 0 & \text{otherwise} \end{cases}$$

where $\epsilon \in \mathfrak{U} \pmod{\mathfrak{P}^m}$. Let us assume now that $\nu \geqq m$. Then

$$\kappa_{11}x(\kappa_{12}x + \kappa_{22})^{-1} = \kappa_{11}\epsilon\tau^l(\kappa_{12}\epsilon\tau^l + \kappa_{22})^{-1},$$

but $\kappa_{12}\epsilon\tau^l + \kappa_{22} \in \mathfrak{U}$ since $l > 0$ and $\kappa_{21} \in \mathfrak{P}^m$ implies $\kappa_{22} \in \mathfrak{U}$. Hence κ maps $\mathfrak{U}\tau^l$ onto itself by linear fractional transformation. Thus

$$\begin{aligned}[u(\kappa)f_l](x) &= \chi[(\kappa_{12}x + \kappa_{22})^{-2} \det \kappa] f_l[\kappa_{11}x(\kappa_{12} + \kappa_{22})^{-1}] \\ &= \chi[(\kappa_{12}\epsilon\tau^l + \kappa_{22})^{-2} \det \kappa]\chi[\kappa_{11}\epsilon(\kappa_{12}\epsilon\tau^l + \kappa_{22})^{-1}]^{-2} \\ &= \chi[(\det \kappa)\kappa_{11}^{-2}\epsilon^{-2}]\end{aligned}$$

which proves that

$$(8.6) \qquad u(\kappa)f_l = \chi(\kappa_{11}^{-2} \det \kappa) f_l.$$

Now if k is given by a matrix (k_{ij}) then $\kappa = y^{-\nu} k y^{\nu}$ means

$$(8.7) \qquad \begin{pmatrix} k_{11}\tau^{\nu} & k_{12} \\ k_{21}\tau^{\nu} & k_{22} \end{pmatrix} = c \begin{pmatrix} \tau^{\nu}\kappa_{11} & \tau^{\nu}\kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{pmatrix}$$

with $c \in \Omega^{\times}$, so that $\kappa_{11}^{-2} \det \kappa = k_{11}^{-2} \det k$. Hence, noting that χ is a character mod $\mathfrak{P}^m$, we see that for $k \in K_{\nu}$

$$(8.8) \qquad u_{\nu}(k) f_l = \chi(k_{11}^{-2} \det k) f_l = \chi(k_{22}/k_{11}) f_l$$

since we assume $\nu \geqq m$ so that $k_{11}^{-2} \det k \equiv k_{22}/k_{11} \pmod{\mathfrak{P}^m}$ because $k_{12} \in \mathfrak{P}^m$ and $k_{11}k_{22} - k_{12}k_{21} \in \mathfrak{U}$.

Now let $wx = -x^{-1}$ as before and set again $\psi_{\infty} = u(w)\psi_0$ where ψ_0 is the characteristic function of $\mathfrak{P}^m$ (multiplied by $q^{m/2}$ as in § 4). Then $u(k)\psi_{\infty} = u(k)\psi_0$ therefore $u(w)^{-1}u(k)\psi_{\infty} = u(w^{-1}kw)\psi_0$. Now $w^{-1}kw$ is given by the matrix

$$\begin{pmatrix} k_{22} & -k_{21} \\ -k_{12} & k_{11} \end{pmatrix} \equiv \begin{pmatrix} k_{22} & -k_{21} \\ 0 & k_{11} \end{pmatrix} \pmod{\mathfrak{P}^m}$$

so that

$$[u(w^{-1}kw)\psi_0](x) = \chi[(-k_{21}x + k_{11})^{-2} k_{11} k_{22}] \psi_0[k_{22}x(k_{11} - k_{21}x)^{-1}]$$

But $\psi_0[k_{22}x(k_{11} - k_{21}x)^{-1}] = 0$ unless $x \in \mathfrak{P}^m$, hence

$$u(w^{-1}kw)\psi_0 = \chi(k_{22}/k_{11})\psi_0$$

i. e.

$$u(kw)\psi_0 = \chi(k_{22}/k_{11}) u(w)\psi_0$$

hence

$$u(k)\psi_{\infty} = \chi(k_{22}/k_{11})\psi_{\infty} \text{ for all } k \in K_{\nu}.$$

Thus if we combine (8.8) and (8.9) we obtain the following result.

LEMMA 8.1. *Assume that* $\nu \geqq m$. *For each* $l = 1, 2, \cdots, m-1$ *the linear transformation, which maps* ψ_{∞} *onto* f_l *and the orthogonal complement of* ψ_{∞} *onto* 0, *is an intertwining operator* X *as defined by* (8.3).

Now we consider the function f_l' defined by

$$(8.10) \qquad f_l' = u(w) f_l.$$

Then for $k \in K_{\nu}$ we have

$$\begin{aligned} [u(w)^{-1}u(k)f_l'](x) &= [u(w^{-1}kw)f_l](x) \\ &= \chi[(-k_{21}x + k_{11})^{-2} \det k] f_l[k_{22}x(k_{11} - k_{21}x)^{-1}]. \end{aligned}$$

Now if $x = \epsilon\tau^l$ then $f_l(x) = \chi(\epsilon)^{-2}$ so that

$$f_l(k_{22}x/k_{11}-k_{21}x)=f_l[k_{22}\epsilon\tau^l(k_{11}-k_{21}\epsilon\tau^l)^{-1}]$$
$$=\chi[k_{22}\epsilon(k_{11}-k_{21}\epsilon\tau^l)^{-1}]^{-2}$$

hence

$$[u(w)^{-1}u(k)f_l'](x)=\chi(k_{22}{}^{-2}\epsilon^{-2}\det k)=\chi(k_{22}{}^{-2}\det k)f_l(x).$$

Therefore

$$u(k)f_l'=\chi(k_{22}{}^{-2}\det k)u(w)f_l$$
$$=\chi(k_{22}{}^{-2}\det k)f_l',$$

but as before $k_{22}{}^{-2}\det k\equiv k_{11}/k_{22}\ (\operatorname{mod}\mathfrak{P}^m)$ and χ is a character $\operatorname{mod}\mathfrak{P}^m$, so that we have proved

$$u(k)f_l'=\chi(k_{11}/k_{22})f_l' \text{ for all } k\in K_\nu. \tag{8.11}$$

Next we determine $u_\nu(k)\psi_0$, i.e. $u(y^{-\nu}ky^{\nu})\psi_0$ $k\in K_\nu$. As before let $\kappa=y^{-\nu}ky^{\nu}$ then we know that κ is given by a matrix (κ_{ij}) with $\kappa_{21}\in\mathfrak{P}^\nu$ so that, using again $\nu\geqq m$, we have

$$[u(\kappa)\psi_0](x)=\chi[(\kappa_{12}x+\kappa_{22})^{-2}\det\kappa_0]\psi_0\left(\frac{\kappa_{11}x}{\kappa_{12}x+\kappa_{22}}\right).$$

which equals 0 unless $x\in\mathfrak{P}^m$. Hence $u(\kappa)\psi_0=\chi(\kappa_{11}/\kappa_{22})\psi_0$. But by (8.7) $\kappa_{11}/\kappa_{22}=k_{11}/k_{22}$ so that

$$u_\nu(k)\psi_0=\chi(k_{11}/k_{22})\psi_0 \text{ for all } k\in K_\nu. \tag{8.12}$$

Thus, if we combine (8.11) with (8.12) we obtain the following result.

LEMMA 8.2. *Assume that $\nu\geqq m$. For each $l=1,2,\cdots,m-1$ the linear transformation X which maps ψ_0 to f_l' and the orthogonal complement of ψ_0 to 0 is an intertwining operator as defined by* (8.3).

Now let us consider the function $h(x)$ defined by

$$h(x)=\begin{cases}1 \text{ if } x\neq\infty \text{ i.e. } x\in\mathfrak{O}\ (\operatorname{mod}\mathfrak{P}^m)\\ 0 \text{ if } x=\infty.\end{cases} \tag{8.13}$$

Again we use $k\in K_\nu$ and $\nu\geqq m$ so that $k_{12}\in\mathfrak{P}^\nu$, hence

$$[u(k)h](x)=\chi(k_{22}{}^{-2}\det k)h[(k_{11}x+k_{21})k_{22}{}^{-1}].$$

But $(k_{11}x+k_{21})k_{22}{}^{-1}\in\mathfrak{O}\ (\operatorname{mod}\mathfrak{P}^m)$ if and only if $x\in\mathfrak{O}\ (\operatorname{mod}\mathfrak{P}^m)$ i.e. $h[(k_{11}x+k_{21})/k_{22}]=h(x)$ so that

$$u(k)h=\chi(k_{11}/k_{22})h \text{ for all } k\in K_\nu. \tag{8.14}$$

Combining (8.12) with (8.14) we obtain

Lemma 8.3. *Let $\nu \geqq m$. Then the linear transformation X which maps ψ_0 onto h and the orthogonal complement of ψ_0 onto* 0 *satisfies* (8.3).

Now we consider the function $h'(x)$ defined by $h' = u(w)h$. We have

$$[u(w^{-1})u(\kappa)h'](x) = [u(w^{-1}\kappa w)h](x)$$
$$= \chi[(\kappa_{11} - \kappa_{21}x)^{-2} \det \kappa] h[\kappa_{22}x(\kappa_{11} - \kappa_{21}x)^{-1}]$$

whenever $\kappa_{12} \in \mathfrak{P}^{\nu}$. But $u(w^{-1}\kappa w)h = \chi(\kappa_{22}/\kappa_{11})h$ by (8.14) so that $u(\kappa)h' = \chi(\kappa_{22}/\kappa_{11})h'$. Setting again $k = y^{\nu}\kappa y^{-\nu}$ and using (8.7) we have as before $\kappa_{22}/\kappa_{11} = k_{22}/k_{11}$ hence for $\nu \geqq m$.

$$u_{\nu}(k)h' = \chi(k_{22}/k_{11})h' \text{ for all } k \in K_{\nu}. \tag{8.15}$$

Thus if we combine (8.9) and 8.15) we obtain

Lemma 8.4. *The linear transformation X which maps ψ_{∞} onto h' and the orthogonal complement of ψ_{∞} onto* 0 *satisfies* (8.3).

Thus, if we combine Lemmas 8.1 to 8.4 we see that there exist at least $2m$ linearly independent matrices X satisfying (8.3). This proves

Theorem 8.1. *Let u be a representation in the principal series* mod $\mathfrak{P}^m$, *considered as a representation of K. Let $\nu \geqq m$. Then there exist at least $2m$ linearly independent elements of S_u which vanish outside the double coset $Ky^{\nu}K$.*

If we combine this with Theorem 7.2, we obtain

Theorem 8.2. *Let $\tilde{u} = \tilde{u}_{\chi}$ be a representation in the principal series* mod $\mathfrak{P}^m$ *corresponding to a primitive congruence character χ* mod q^m. *Assume that q is not a power of 2. Let u be the corresponding representation of K (see* § 3 *especially* (3.3) *and* (3.5)).

If $m > 1$ then $\mathfrak{S}_u^0 \neq S_u^0$, in fact $\mathfrak{S}_u(\nu) \neq S_u(\nu)$ for $\nu \geqq m$.

However when $m = 1$ we can show that $\mathfrak{S}_u = S_u$. Indeed we have

Theorem 8.3. *Let $u = u_{\chi}$ be an irreducible representation of K belonging to the principal series* mod $\mathfrak{P}$, *corresponding to a non-real congruence character χ* mod $\mathfrak{P}$. *Assume that q is not a power of 2. Then $\mathfrak{S}_u = S_u$.*

Proof. We consider again the restriction of u to the subgroup K_{ν}. Since $k \in K_{\nu}$ means $k_{12} \in \mathfrak{P}^{\nu}$ and u is now in the principal series mod $\mathfrak{P}$, we have for $k \in K_{\nu}$

$$[u(k)\Psi](x) = \chi(k_{22}^{-2} \det k)\Psi\left(\frac{k_{11}x + k_{21}}{k_{22}}\right). \tag{8.16}$$

Hence the constant multiples ch of h form an invariant ray under $u(K_\nu)$ and so do the constant multiples $c\Psi_\infty$ of Ψ_∞. Let $\mathfrak{v}'$ denote the orthogonal complement of the sum of these two rays in $\mathfrak{v}$. Then $\mathfrak{v}'$ is a $u(K_\nu)$-invariant subspace of $\mathfrak{v}$ of dimension $q-1$. It is easy to see that $\mathfrak{v}'$ is irreducible under $u(K_\nu)$. Similarly if $k \in K_\nu{}^t$ then $\{c\Psi_0\}$ and $\{ch'\}$ are two rays, each invariant under $u(k)$. Let $\mathfrak{v}''$ be the orthogonal complement in $\mathfrak{v}$ to these two rays. Again one sees that $\mathfrak{v}''$ is irreducible under $u(K_\nu{}^t)$. We have to consider these two irreducible representations of K_ν: The restriction of $u(K_\nu)$ to $\mathfrak{v}'$ and the restriction of $u_\nu(K_\nu)$ to $\mathfrak{v}''$. These two representations of K_ν are inequivalent. Indeed let k be given by the matrix $\begin{pmatrix} 1 & \tau^\nu \\ 0 & 1 \end{pmatrix}$. Then $k^{(\nu)}$ is given by the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Since $\nu \geqq 1$ and u is a representation mod $\mathfrak{P}$, we have $u(k) = I$. On the other hand $u(k^{(\nu)}) \neq I$ as is seen at once, using the explicit form (8.16) of u as multiplier representation. But $u_\nu(k) = u(k^{(\nu)})$, hence $\mathfrak{v}'$ and $\mathfrak{v}''$ are inequivalent irreducible representation spaces of K_ν for each $\nu = 1, 2, \cdots$.

Hence the only possible intertwining operators X satisfying (8.3) are those between the rays $\{ch\}$, $\{ch'\}$, $\{c\Psi_0\}$ and $\{c\Psi_\infty\}$. By Lemmas 8.1 to 8.4 (see also (8.12), (8.14) and (8.15)) there are exactly two linearly independent ones in the present case $m = 1$. Using now Theorem 7.2 we have thus proved Theorem 8.3.

9. Discrete series. We wish to construct irreducible unitary representations of G whose matrix-coefficients (with respect to suitable orthonormal bases) are square integrable functions on G. For this purpose we start with an irreducible unitary representation u of K, whose restriction to the subgroup of translations does not contain the identity representation, i. e.

$$u\left[\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\right] \not\ni I. \tag{9.1}$$

We consider again the algebra $S_u{}^0$. If $\mathfrak{f}(g)$ is contained in $S_u{}^0$ then we have as above that $\mathfrak{f}(ky^\nu) = \mathfrak{f}(y^\nu\mathfrak{k})$ whenever $ky^\nu = y^\nu\mathfrak{k}$ and y is given by the matrix $\begin{pmatrix} \tau & 0 \\ 0 & 1 \end{pmatrix}$. Letting $X = \mathfrak{f}(y^\nu)$ for fixed $\nu \geqq 1$ we have as before

$$u(k)X = Xu_\nu(k) \text{ for all } k \in K_\nu \tag{9.2}$$

(see § 2 above). Now let $k = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ then $k^{(\nu)} = y^{-\nu}ky^\nu = \begin{pmatrix} 1 & 0 \\ \tau^\nu & 1 \end{pmatrix}$. But u is a continuous representation of K, hence there exists a (smallest) positive integer m such that $u(K^{(m)}) = I$, hence

$$u\left[\begin{pmatrix}1 & 0\\ \tau^\nu & 1\end{pmatrix}\right]=I \text{ for } \nu\geqq m. \tag{9.3}$$

Thus combining (9.2) and (9.3) we obtain for $k=\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$ and $\nu\geqq m$

$$u\left[\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}\right] X=X. \tag{9.4}$$

But if we now use (9.1) we see that (9.4) and hence also (9.2) has no solutions X other than $X=0$. This proves

LEMMA 9.1. *Let u be an irreducible unitary representation of K whose restriction to the subgroup of translations does not contain the one-representation, i.e.* (9.1) *holds. Assume $u(K^{(m)})=I$. Then $\mathfrak{f}(g)\in S_u$ implies that $\mathfrak{f}(g)$ vanishes on $Ky^\nu K$ whenever $\nu\geqq m$. Hence in particular S_u is a finite dimensional vector space over the complex numbers.*

It follows (see for example Mautner [4]) that the commuting algebra of the induced representation $\operatorname*{ind}_{K\uparrow G} u$ is a finite dimensional $*$-algebra of bounded operators on the given Hilbert space $\mathfrak{H}$. Hence $\operatorname{ind} u$ decomposes into a direct sum of a finite number of irreducible representations, say $U^{(j)}$, of G:

$$\mathfrak{H}=\sum_{j=1}^{\oplus}\mathfrak{H}^{(j)},\ \operatorname*{ind}_{K\uparrow G} u=\sum_{j=1}^{\oplus} U^{(j)}. \tag{9.5}$$

Let us note that if $m=1$, then $\mathfrak{f}(g)$ vanishes on $Ky^\nu K$ for every $\nu\geqq 1$, i.e. $f(g)$ vanishes outside K and $\mathfrak{f}(k)=cu(k)$ so that S_u consists of the scalar multiples of the identity only; thus $\operatorname{ind} u$ is irreducible in this case.

In any case let F_j be the orthogonal projection of $\mathfrak{H}$ onto $\mathfrak{H}^{(j)}$ and $\mathfrak{f}_j(g)$ the corresponding function $\in S_u{}^0$. Let $v(g)$ be a column of $\mathfrak{f}_j(g)$ considered as a vector-valued function on G; clearly $v(kg)=u(k)v(g)$ and $v(g)$ is a continuous function of compact support satisfying $\mathfrak{f}_j * v=v$ so that $v\in\mathfrak{H}^{(j)}$. Assuming $\mathfrak{H}^{(j)}\neq(0)$ we have $\mathfrak{f}_j(g)$ not identically zero, so that v can be chosen $\neq 0$. Now $U^{(j)}$ is irreducible and $[U^{(j)}(g^1)v](g)=v(gg^1)$, so that the finite linear combinations of all right translates of v are dense in $\mathfrak{H}^{(j)}$. Let us choose a countable dense subset of them and apply the Schmidt orthogonalization process: We obtain a complete orthonormal sequence $v_n\in H^{(j)}$ such that each $v_n(g)$ is of compact support and continuous. Hence

$$\langle U^{(j)}(g)v_m, v_n\rangle=\int_G\langle v_m(xg), v_n(x)\rangle\,dx$$

is for each m, n of compact support on G. We have thus proved the following result.

THEOREM 9.1. *Let u be an irreducible unitary representation of K whose restriction to the translations does not contain the one-representation. Let U be the induced representation:*

$$U = \operatorname*{ind}_{K \uparrow G} u$$

Then U is the direct sum of a finite number of irreducible unitary representations $U^{(j)}$ of G. In a suitable orthonormal base the matrix coefficients of each $U^{(j)}$ are functions of compact support on G.

Let us note that representations u of K with the above property (9.1) exist, as was already known to Frobenius [1] and Schur [1].

REFERENCES.

F. Bruhat.

1. "Distributions sur un groupe localement compact et applications à l'étude des representations des groupes p-adiques," *Bulletin de la Société Mathématique de France,* vol. 89 (1961), pp. 43-74.
2. "Sur les répresentations des groupes classiques p-adiques I and II," *American Journal of Mathematics,* vol. 83 (1961), pp. 321-338 and 343-368.

G. Frobenius.

1. "Ueber die Darstellung der endlichen Gruppen," *Sitzungsberichte der Deutschen Akademie der Wissenschaften zu Berlin* (1896), pp. 994-1015.

G. W. Mackey.

1. "Induced representations of locally compact groups I and II," *Annals of Mathematics,* vol. 55 (1952), pp. 101-139 and vol. 58 (1953), pp. 193-221.

F. I. Mautner.

1. "Spherical functions over $\mathfrak{p}$-adic fields, I," *American Journal of Mathematics,* vol 80 (1958), pp. 441-457.
2. "Fourier analysis and symmetric spaces," *Proceedings of the National Academy of Sciences, U. S. A.,* vol. 37 (1951), pp. 529-533.
3. "Note on the Fourier inversion formula on groups," *Transactions of the American Mathematical Society,* vol. 78 (1955), pp. 371-384.
4. "Induced representations," *American Journal of Mathematics,* vol. 84 (1952), pp. 737-758.

I. Schur.

1. "Untersuchungen über die Darstellungen der endlichen Gruppen," *Journal für die Reine und Angewandte Mathematik,* vol. 132 (1907), pp. 85-137.

COMPACTIFICATIONS OF SYMMETRIC SPACES.*

By Calvin C. Moore.

Section 1. Furstenberg in a recent work [2] has introduced the notions of a boundary and of boundary subgroups of a connected Lie group. If G is a semi-simple group with finite center and K a maximal compact subgroup, then these boundaries can be used to compactify the Riemannian symmetric space G/K. In this paper we shall study in detail the structure of these boundaries and the resulting compactifications. Also we shall show that this construction leads to essentially the same compactifications which Satake [8] has studied. Some of our results are in Satake's paper, but we need them in modified forms. Hermann has also studied these boundaries and has obtained many interesting results, [4, 5], using more geometrical methods. There is some overlap in the results.

We shall briefly recall Furstenberg's results and establish our conventions. Let $M = G/H = \{xH\,;\, x \in G\}$ be a homogeneous space of the separable locally compact group G. If μ is a Borel measure on M and $g \in G$, we define the translate $g \cdot \mu$ of μ by the formula $g \cdot \mu(E) = \mu(g^{-1} \cdot E)$ for a Borel set E. Then Furstenberg calls M a boundary of G and (H a boundary subgroup) if M is compact, and if for any probability measure μ on M, we can find a sequence $g_n \in G$ such that $g_n \cdot \mu$ tends in the weak-* topology to a point mass. (A sequence μ_n converges to μ in the weak-* topology if and only if

$$\int f\, d\mu_n \to \int f\, d\mu$$

for any continuous function on M.) We observe, [7] that on M there is a unique equivalence class of measures (called quasi-invariant measures) whose null sets are invariant under translation by G. Then Furstenberg has shown that for M to be a boundary it is enough that M be compact and that there is a quasi-invariant probability measure μ on M and a sequence g_n in G such that $g_n \cdot \mu$ tends to a point mass on M in the weak-* topology.

Further Furstenberg shows that there exists minimal boundary subgroups and all such are conjugate. Moreover using the Iwasawa structure theory [6] one can see that there is a normal subgroup L of G such that L is contained

---

Received July 11, 1963.

* Research supported in part by National Science Foundation Grant 19874.

in a minimal boundary subgroup H of G (and hence all) such that G/L is a semi-simple Lie group with finite center. Since L acts trivially on G/H, it is clear that all questions of structure can be reduced to the case when $L = (0)$. We shall thus treat just this case.

If G is a semi-simple linear algebraic group, then Theorem 1 implies that the boundary subgroups of G are exactly the parabolic subgroups in the sense of Borel-Godement [10]. Some of our results can be obtained using this theory.

In Section 2 we shall compute the structure of minimal boundary subgroups. Sections 3 and 4 extend these results to all boundary subgroups. In Sections 5 and 6 we shall study the resulting compactifications of the symmetric space G/K where K is a maximal compact subgroup of G. Finally in Section 7 we shall examine the connection with Satake's compactifications [8].

We are greatly indebted to R. Hermann, H. Furstenberg, A. Koranyi and D. Wehn for many interesting and useful discussions of these topics.

Section 2. In this section we shall compute the structure of the minimal boundary subgroups of G in terms of the known structure of G (a semi-simple group with finite center). For the following see for instance [9]. Let $\mathfrak{g}_0$ be the Lie algebra of G, $\mathfrak{g}$ its complexification, $\mathfrak{k}_0$ a maximal compact subalgebra of $\mathfrak{g}_0$ and $\mathfrak{p}_0$ the orthogonal complement of k_0 in $\mathfrak{g}_0$ with respect to the Killing form. We can choose a compact form $\mathfrak{u}$ of $\mathfrak{g}$ such that $\mathfrak{u} = \mathfrak{k}_0 + i\mathfrak{p}_0$ $(i = \sqrt{-1})$, and $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$. Let σ be the involution of $\mathfrak{g}$ having $\mathfrak{g}_0$ as fixed points, and let $\mathfrak{h}^-$ be a maximal abelian subspace of $\mathfrak{p}_0$; then we can find a subspace $i\mathfrak{h}^+ \subset \mathfrak{k}_0$ such that $\mathfrak{h}_0 = \mathfrak{h}^- + i\mathfrak{h}^+$ is a Cartan subalgebra of $\mathfrak{g}_0$. Let Σ be the set of roots of $\mathfrak{g}$ with respect to the complexification $\mathfrak{h}$ of $\mathfrak{h}_0$. The real subspace spanned by the roots is $\mathfrak{h}^- + \mathfrak{h}^+$. We choose a basis for $\mathfrak{h}^- + \mathfrak{h}^+$ by composing a basis in $\mathfrak{h}^-$ and one in $\mathfrak{h}^+$, and we fix once and for all the lexicographic order on Σ. The conjugate linear automorphism σ preserves $\mathfrak{h}$ and so permutes roots by $(\sigma(\alpha)(H)) = \overline{\alpha(\sigma(H))}$ for $\alpha \in \Sigma$, $H \in \mathfrak{h}$. Now Σ can be decomposed into three disjoint sets; $\Sigma^+ = \{\alpha : \alpha, \sigma(\alpha) > 0\}$, $\Sigma^- = -\Sigma^+$, and $\Delta = \{\alpha : \alpha = -\sigma(\alpha)\}$; it is clear that Σ^+ (Σ^-) consists of those positive (negative) roots which do not vanish on $\mathfrak{h}^-$, and Δ consists of those roots vanishing on $\mathfrak{h}^-$. We write

$$\mathfrak{n} = \Sigma_\alpha C \cdot X_\alpha \qquad (\alpha \in \Sigma^+)$$
$$\mathfrak{n}^- = \Sigma_\alpha C \cdot X_\alpha \qquad (\alpha \in \Sigma^-)$$

where X_α are root vectors. Then $\mathfrak{n}$ and $\mathfrak{n}^-$ are nilpotent subalgebras invariant under σ. We denote by $\mathfrak{n}_0$ and $\mathfrak{n}_0^-$ their intersections with $\mathfrak{g}_0$. Then the following are known, [1] 1°. $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{h}^- + \mathfrak{n}_0$ as direct sum, and $\mathfrak{l}_0 = \mathfrak{h}^- + \mathfrak{n}_0$ is a solvable subalgebra.

2°. If $\mathfrak{m} = \mathfrak{h}^+ + i\mathfrak{h}^+ + \Sigma \boldsymbol{C} \cdot X_\alpha$ $(\alpha \in \Delta)$, then $\mathfrak{m}$ is a subalgebra of $\mathfrak{k}$ invariant under σ. Also $\mathfrak{m}(\mathfrak{m}_0 = \mathfrak{m} \cap \mathfrak{g}_0)$ is the centralizer and normalizer of $\mathfrak{h}^-$ in $\mathfrak{k}(\mathfrak{k}_0)$.

3°. $\mathfrak{b}_0 = \mathfrak{m}_0 + \mathfrak{h}^- + \mathfrak{n}_0$ is a subalgebra normalizing $\mathfrak{n}_0$ and $\mathfrak{l}_0$.

Finally, let K, A, N, N^-, L, M_0 be the analytic subgroups of G corresponding to $\mathfrak{k}_0$, $\mathfrak{h}^-$, $\mathfrak{n}_0$, $\mathfrak{n}_0^-$, $\mathfrak{l}_0$, and $\mathfrak{m}_0$. They are all closed and $G = KAN$ is an Iwasawa decomposition. Let M and $\hat{M}$ be respectively the centralizer and normalizer of $\mathfrak{h}^-$ in K. Then M_0 is the connected component of both M and $\hat{M}$.

The following Lemma is known, [1] and [8].

Lemma 1. *Every coset of M_0 in $\hat{M}$ contains an element x such that* $\operatorname{ad}(x)(i\mathfrak{h}^+) = i\mathfrak{h}^+$. *Moreover $M = M_0 \cdot Z$ where $Z = \exp(i\mathfrak{h}^-) \cap K$ is a finite abelian group central in M.*

For the following, see [1] and [3].

Lemma 2. *M normalizes N and $B = MAN$ is a closed subgroup normalizing N and L.*

The main result of this section has been conjectured by Furstenberg.

Theorem 1. *B is the normalizer of N and L and is a minimal boundary subgroup.*

Proof. Let B' be the normalizer of N. Then $B' \supset B$ and by Bruhat's double coset lemma, [1], and Lemma 1 any $x \in B'$ can be written $x = b'yb$ with $y \in M$ such that $\operatorname{ad}(y)(ih^+) = (ih^+)$, and with $b', b \in B$. It follows that $\operatorname{ad}(y)$ permutes the elements of Σ and since y normalizes N, $\operatorname{ad}(y)$ permutes the roots of Σ^+ among themselves. If Λ denotes the open positive Weyl chamber in $\mathfrak{h}^-$ (i.e. $\{H : \alpha(H) > 0, \alpha \in \Sigma^+\}$), then $\operatorname{ad}(y)\Lambda = \Lambda$. Now $\operatorname{ad}(y)$ as linear transformation on $\mathfrak{h}^-$ has finite order since $\hat{M}/M$ is finite. Therefore since Λ is convex, it is easy to see that $\operatorname{ad}(y)$ has a fixed point in Λ. But all the elements of Λ are regular ([1], p. 188), and if y centralizes a regular element of $\mathfrak{h}^-$, it centralizes all of $\mathfrak{h}^-$. Thus $y \in M$ and $x \in B$, and so $B = B'$ as desired. Finally if x normalizes L, it normalizes N since N is the derived group of L. Thus $x \in B' = B$, but it is clear that B normalizes L and the first statement is proved.

We now turn to the second part of the theorem. Furstenberg has shown that the normalizer of L, which is B, is contained in a minimal boundary subgroup. Consequently, all we need show is that G/B is a boundary.

By the Bruhat double coset lemma, each $B \cdot x \cdot B$ meets $\hat{M}$ in a coset of M, and $B\backslash G/B$ may be identified with $\hat{M}/M$. Let $s_1, \cdots, s_n$ be elements of $\hat{M}$ representing the cosets of M such that $\mathrm{ad}(s_i)\mathfrak{h}_0 = \mathfrak{h}_0$. Let $s'_i = s_iB \in G/B$. Then the orbit of B in G/B through s'_i is

$$B \cdot s'_i \cong B/s_iBs_i^{-1} \cap B = MAN/MA(\mathrm{ad}(s_i)N \cap N).$$

Since $\mathrm{ad}(s_i)$ permutes the roots of $\Sigma^+ \cup \Sigma^-$, $\mathrm{ad}(s_i)\mathfrak{n}_0 \subset \mathfrak{n}_0 + \mathfrak{n}_0^-$. Now there are a finite number (n) of orbits of B on G/B. Consequently at least one of them must have dimensional equal to that of G/B. But we know that $\dim(G/B) = \dim N^- = \dim N$, and $\dim B \cdot s'_i = \dim N - \dim(\mathrm{ad}(s_i)N \cap N)$. It follows that the orbits $B \cdot \hat{s}_i$ of maximal dimension are characterized by $\mathrm{ad}(s_i)\mathfrak{n}_0 \cap \mathfrak{n}_0 = (0)$, or equivalently that $\mathrm{ad}(s_i)\mathfrak{n}_0 = \mathfrak{n}_0^-$. Suppose now that s_i and s_j represent two orbits of maximal dimension. Then $\mathrm{ad}(s_is_j^{-1})\mathfrak{n}_0 = \mathfrak{n}_0$, and $s_is_j^{-1}$ normalizes $\mathfrak{n}_0$. Thus $s_is_j^{-1} \in B$ and $s_i = s_j$. Therefore we find that there is exactly one orbit of highest dimension, $B \cdot \hat{s}_0$. Now the complement of $B \cdot s'_0$ is a finite union of submanifolds of lower dimension and hence a set of quasi-invariant measure 0. Let $P = s_0^{-1}B \cdot s'_0$; then $G/B - P$ has quasi-invariant measure 0. We claim that P is the one-one image in G/B of the subgroup N^- of G. By the above, every element of $B \cdot s'_0$ can be written uniquely in the form ns_0B for some $n \in N$. Thus any element of P can be written uniquely as $s_0^{-1}ns_0B = \mathrm{ad}(s_0^{-1})(n)B$. But $\mathrm{ad}(s_0^{-1})$ maps N homeomorphically onto N^- and our assertion is proved. We now identify P with $\mathfrak{n}_0^-$ by $n \to \exp(n)B$ for $n \in \mathfrak{n}_0^-$. Moreover if $a \in A$ and $n \in \mathfrak{n}_0^-$,

$$a \cdot \exp(n)B = \exp(\mathrm{ad}(a)n)a^{-1}B = \exp(\mathrm{ad}(a)n)B.$$

Thus A leaves P invariant and acts on P just as A acts on $\mathfrak{n}_0^-$ by the adjoint representation. Moreover $\mathrm{ad}(a)$ for $a \in A$ on the complexification $\mathfrak{n}^-$ of $\mathfrak{n}_0^-$ is diagonalizable with eigenvalues $\exp(\alpha(\log a))$ for $\alpha \in \Sigma^-$. (We use $\log(a)$ to denote the unique element of $\mathfrak{h}^-$ such that $\exp(\log(a)) = a$.) These eigenvalues are real and it follows that $\mathrm{ad}(a)$ on $\mathfrak{n}_0^-$ is diagonalible with the same eigenvalues. We introduce coordinates $y_\alpha(\alpha \in \Sigma^-)$ on $\mathfrak{n}_0^-$ and corresponding coordinates on P which effect this diagonalization.

Now let $H \in \Lambda = \{H\colon \alpha(H) > 0, \alpha \in \Sigma^+\}$, and let $a_t = \exp(tH)$, and let μ be a quasi-invariant probability measure on G/B. We claim that $a_t \cdot \mu$ tends to a point mass at B. Let f be a continuous function on G/B with compact support such that $0 \leqq f \leqq 1$ and $f = 1$ on S_b where $S_b = \{p\colon p \in P$ and $|y_\alpha(p)| \leqq b\}$. Then

$$\int f\, d(a_t\mu) = \int_P f(a_t x)\, d\mu(x)$$

since $\mu(P) = 1$. We now see that

$$\mu(a_t^{-1} S_b) \leqq \int_P f(a_t x)\, d\mu(x) \leqq 1.$$

We may find $d > 0$ so that $\alpha(H) > d$ for all $\alpha \in \Sigma^+$, and consequently the eigenvalues of a_t^{-1} on $\mathfrak{n}_0^-$ all exceed $\exp(td)$, and $a_t^{-1} S_b \supset S_{b\exp(td)}$. Thus as $t \to \infty$,

$$1 \geqq \mu(a_t^{-1} S_b) \geqq \mu(S_{b\exp(td)}) \to \mu(P) = 1,$$

and

$$\int f\, d(a_t\mu) \to 1.$$

Now we may choose by the weak-* compactness of the set of probability measures on G/B, a subsequence t_n such that $a_{t_n} \cdot \mu$ converges to a limit μ'. It follows that

$$\int f\, d\mu' = 1.$$

Since we can choose f to vanish outside S_{2b}, it is clear that μ' is a point mass at the origin B in G/B. This holds for the limit of any subsequence $a_{s_n} \cdot \mu$ which converges at all. It follows at once that $a_t \cdot \mu$ tends to a point mass at B. Our theorem is proved.

Section 3. It now follows that any closed subgroup C of G which contains B is also a boundary subgroup. We shall classify such subgroups in this section. Let $\alpha_1, \cdots, \alpha_n$ be a fundamental system of roots of $\mathfrak{g}$ with respect to our fixed order, and let π denote the restriction map of linear forms on $\mathfrak{h}_0$ to $\mathfrak{h}^-$. Then we denote by F the distinct non-zero vectors among the set $\{\pi(\alpha_i)\}$. Following Satake, F is called a restricted fundamental system for $\mathfrak{g}_0$. The vectors of F are linearly independent [8] and if $\alpha \in \Sigma^+, \pi(\alpha) = \sum n_\gamma \gamma$ $(\gamma \in F)$ where the n_γ are non-negative integers.

Following Satake, we make the following definitions. If E is a subset of F, let $\Delta(E) = \{\alpha \in \Sigma : \pi(\alpha) = \sum n_\gamma \gamma \ (\gamma \in E)\}$ and let $\Sigma^+(E)$ $(\Sigma^-(E))$ be the remaining positive (negative) roots. Let $\mathfrak{h}'^-(E) = \{H \in \mathfrak{h}^- : \gamma(H) = 0, \gamma \in E\}$ and let $\mathfrak{h}^-(E)$ be the orthogonal complement of $\mathfrak{h}'^-(E)$ in $\mathfrak{h}^-$. Finally let $\mathfrak{h}(E)$ be the complex vector space spanned by the H_α with $\alpha \in \Delta(E)$. Since $\sigma(\Delta(E)) = \Delta(E)$, $\mathfrak{h}(E)$ is the complexification of $\mathfrak{h}_0(E) = \mathfrak{h}(E) \cap \mathfrak{g}_0$. Now it is easy to see that $\mathfrak{h}_0(E) = i\mathfrak{h}^+(E) + \mathfrak{h}^-(E)$ where $i\mathfrak{h}^+(E) = \mathfrak{h}_0(E) \cap i\mathfrak{h}^+$. We let

$$\mathfrak{g}(E) = \mathfrak{h}(E) + \sum \boldsymbol{C} \cdot X_\alpha \quad (\alpha \in \Delta(E))$$

$$\mathfrak{n}(E) = \sum \boldsymbol{C}\cdot X_\alpha \quad (\alpha \in \Sigma^+(E)) \text{ and similarly for } \mathfrak{n}^-(E).$$

Further let

$$\mathfrak{l}(E) = (\mathfrak{h} \ominus \mathfrak{h}(E)) + \mathfrak{n}(E) \text{ and}$$

$$\mathfrak{b}(E) = \mathfrak{g}(E) + \mathfrak{l}(E), \text{ where } \ominus \text{ denotes orthogonal complement.}$$

Since σ leaves $\Delta(E)$ and $\Sigma^{\pm}(E)$ invariant, these subalgebras are spanned by their intersections with $\mathfrak{g}_0$, which we denote by $\mathfrak{g}_0(E)$ and so on. Notice that $\mathfrak{g}(E)$ is semi-simple and that $\mathfrak{l}(E)$ is the radical of $\mathfrak{b}(E)$.

Theorem 2. *$\mathfrak{b}_0(E)$ is the normalizer of $\mathfrak{n}_0(E)$ and $\mathfrak{l}_0(E)$, and is self-normalizing. Each $\mathfrak{b}_0(E) \supset \mathfrak{b}_0$ and every subalgebra of $\mathfrak{g}_0$ containing $\mathfrak{b}_0$ is of the form $\mathfrak{b}_0(E)$.*

Proof. It suffices to prove the first statement for the complexifications of these algebras. First we observe that $\mathfrak{h}$ normalizes $\mathfrak{n}(E)$ and $\mathfrak{l}(E)$. Now if $\alpha \in \Delta(E)$, $\beta \in \Sigma^+(E)$ and if $\alpha + \beta$ is a root, it must belong to $\Sigma^+(E)$. Thus $\mathfrak{g}(E)$ and hence $\mathfrak{l}(E)$ normalize $\mathfrak{n}(E)$. If $H \in (\mathfrak{h} \ominus \mathfrak{h}(E))$ and $\alpha \in \Delta(E)$, then $\alpha(H) = 0$. Thus $\mathfrak{g}(E)$ and $\mathfrak{b}(E)$ normalize $\mathfrak{l}(E)$. Since $\mathfrak{g} = \mathfrak{n}^-(E) + \mathfrak{b}(E)$, it suffices to show that no non-zero element of $\mathfrak{n}^-(E)$ normalizes $\mathfrak{n}(E)$ or $\mathfrak{l}(E)$. Let $X = \sum a_\beta X_\beta$, $\beta \in \Sigma^-(E)$ with $a_\alpha \neq 0$. We can find $H \in \mathfrak{h} \ominus \mathfrak{h}(E) \subset \mathfrak{l}(E)$ such that $\alpha(H) \neq 0$. Then

$$[H, X] = a_\alpha \alpha(H) X_\alpha + \sum a_\gamma X_\gamma$$

and this is clearly not in $\mathfrak{l}(E)$ and X does not normalize $\mathfrak{l}(E)$. Also $X_{-\alpha} \in \mathfrak{n}(E)$ and by inspection $[X_{-\alpha}, X]$ is not in $\mathfrak{n}(E)$ so X does not normalize $\mathfrak{n}(E)$. Finally if X normalizes $\mathfrak{b}(E)$, it normalizes $\mathfrak{l}(E)$, the radical of $\mathfrak{b}(E)$, and hence $X \in \mathfrak{b}(E)$.

Now let $\mathfrak{c}_0$ be a subalgebra of $\mathfrak{g}_0$ containing $\mathfrak{b}_0$ and let $\mathfrak{c}$ be its complexification. Since $\mathfrak{c} \supset \mathfrak{h}$ $\mathfrak{c} = \mathfrak{h} + \sum \boldsymbol{C}\cdot X_\alpha$ $(\alpha \in S)$ where S is some set of roots. We shall show that $S = \Delta(E) \cup \Sigma^+(E)$ for some E. In fact let $E = \{\gamma, \gamma \in F \text{ and } \exists\, X_\alpha \in \mathfrak{c} : \pi(\alpha) = -\gamma\}$. Since $S \supset \Delta \cup \Sigma^+$ in virtue of the fact that $\mathfrak{c} \supset \mathfrak{b}$, it suffices to show that $S \cap \Sigma^- = \Delta(E) \cap \Sigma^-$. Let $\alpha \in \Delta(E)$, then $\alpha = -\Sigma\, n_i\alpha_i$ and if $n_i > 0$, then $\pi(\alpha_i) \in E \cup \{0\}$. If $\pi(\alpha_i) = 0$, $X_{-\alpha_i} \in \mathfrak{b} \subset \mathfrak{c}$. If $\pi(\alpha_i) = \gamma \in E$, we can find a root β such that $X_\beta \in \mathfrak{c}$ and $\pi(\beta) = -\gamma$. Then it follows from [8], p. 80 that $+\beta = -\alpha_k - \Sigma\, n_j\alpha_j$ $(\alpha_j \in \Delta)$ where $\pi(\alpha_k) = \gamma$. But now $X_{-\alpha_k}$ is in the algebra generated by X_β, X_{α_k} and X_{α_j} $(\alpha_j \in \Delta)$. Thus $X_{-\alpha_k}$ is in $\mathfrak{c}$. Now $\pi(\alpha_k) = \pi(\alpha_i) = \gamma$ and it follows ([8], p. 80) that either $\alpha_k = \alpha_i$ in which case $X_{-\alpha_i} \in \mathfrak{c}$, or that $\sigma(\alpha_i) = \alpha_k + \Sigma\, n_l\alpha_l$ $(\alpha_l \in \Delta)$. We can argue as before that $X_{-\sigma(\alpha_i)} \in \mathfrak{c}$. But $\mathfrak{c}$ is invariant under

σ, so that $X_{-\alpha_i} \in \mathfrak{c}$. We have now shown that if $\alpha \in \Delta(E)$, and $\alpha = -\Sigma\, n_i\alpha_i$, then $n_i > 0$ implies that $X_{-\alpha_i} \in \mathfrak{c}$. But X_α is in the algebra generated by those $X_{-\alpha_i}$ with $n_i > 0$. Thus $X_\alpha \in \mathfrak{c}$ and $\alpha \in S$. Conversely, let $\alpha \in S \cap \Sigma^-$. Then $\alpha = -\Sigma\, n_i\alpha_i$, and if $n_i > 0$, $X_{-\alpha_i}$ is in the algebra generated by X_α and the X_{α_k}. It follows that if $n_i > 0$, $X_{-\alpha_i} \in \mathfrak{c}$. Then $\pi(\alpha_i) \in E \cup \{0\}$, and it is clear that $\alpha \in \Delta(E)$. Therefore $S \cap \Sigma^- = \Delta(E) \cap \Sigma^-$, and Theorem 2 is proved.

If $\mathfrak{k}_0(E) = \mathfrak{g}_0(E) \cap \mathfrak{k}_0$, then $\mathfrak{k}_0(E)$ is a maximal compact subalgebra of $\mathfrak{g}_0(E)$ which also centralizes $\mathfrak{h}'_0{}^-(E)$. ([8], p. 83.) Let $\mathfrak{m}_0(E)$ be the centralizer of $\mathfrak{h}'_0{}^-(E)$ in $\mathfrak{k}_0$.

Let $B(E)$ be the normalizer of $\mathfrak{n}_0(E)$ in G, and let $M(E)$ be the centralizer of $\mathfrak{h}'_0{}^-(E)$ in K. The connected component $B_0(E)$ of $B(E)$ is the analytic group of G corresponding to $\mathfrak{b}_0(E)$. Finally let $G(E)$ and $K(E)$ be the analytic subgroups corresponding to $\mathfrak{g}_0(E)$, and $\mathfrak{k}_0(E)$. These are both closed, [8]. Part of the following theorem is contained in [8], pp. 91-92.

THEOREM 3. *All the subgroups $B(E)$ contain B, and any closed subgroup of G containing B is of the form $B(E)$. Moreover*

$$B(E) = M(E) \cdot A \cdot N$$

and

$$B(E) = G(E) \cdot Z \cdot L(E)$$

where Z is given in Lemma 2.

Proof. It is clear that B normalizes $\mathfrak{n}_0(E)$ since $B = B_0 \cdot Z$ and B_0 and Z normalize $\mathfrak{n}_0(E)$. Let C be a closed subgroup containing B. Then the Lie algebra $\mathfrak{c}_0$ of C is equal to $\mathfrak{b}_0(E)$, and thus $C \supset B_0(E)$. Now it will follow from the last statement of the theorem that $B(E) = B_0(E) \cdot Z$ since $B_0(E) = G(E) \cdot L(E)$ and $Z \cdot L(E) = L(E) \cdot Z$. But since $C \supset B \supset Z$, it is clear that $C \supset B(E)$. If $x \in C$, $\mathrm{ad}(x)$ must preserve $\mathfrak{c}_0 = \mathfrak{b}_0(E)$ and hence also $\mathfrak{n}_0(E)$ which is the derived algebra of the radical of $\mathfrak{b}_0(E)$. Thus $x \in B(E)$ and $C = B(E)$.

It suffices now to establish the decomposition assertions for $B(E)$. Suppose now that an element $s_i \in M$ normalizes $\mathfrak{n}_0(E)$; we claim that it must be in $M(E)$. This is clear if $E = F$. In general let $\Lambda'(E)$ be the positive Weyl chamber in $\mathfrak{h}'^-(E)$; that is $\{H : \alpha(H) > 0 \text{ for } \alpha \in \Sigma^+(E)\}$. Now $\mathrm{ad}(s_i)$ permutes the roots of $\Sigma^+(E)$ since it normalizes $\mathfrak{n}_0(E)$, and hence $\mathrm{ad}(s_i)\Lambda'(E) = \Lambda'(E)$. We may argue as before that $\mathrm{ad}(s_i)$ has a fixed point in $\Lambda'(E)$ and hence that s_i centralizes $\mathfrak{h}'^-(E)$.

We now claim that $B(E) \subset M(E)A \cdot N$. Let $x \in B(E)$, then $x = bsb'$ with $b, b' \in B$ and $s \in \hat{M} \cap B(E)$ by Bruhat's double coset lemma. By the previous paragraph, $s \in M(E)$, and we write $x = s(s^{-1}bs)b'$. Since $B = MAN$, $b = n \cdot a \cdot m$ where $a \in A$, $m \in M$ and $n \in N$. Further $n = n_1 n_2$ where $n_2 \in N(E)$ and $n_1 \in \exp(\mathfrak{n}'_0(E))$ where $\mathfrak{n}'_0(E) = \mathfrak{n}_0 \cap \mathfrak{g}_0(E)$. (This is possible since $\mathfrak{n}_0(E)$ is an ideal in $\mathfrak{n}_0$ and $\mathfrak{n}'_0(E)$ is a complementary subalgebra.) Then $x = s(s^{-1}n_1 s) \cdot s^{-1}(n_2 am)s \cdot b'$, and it suffices to prove that $s^{-1}n_1 s \in M(E) \cdot A \cdot N$ since $s \in M(E)$ and the final term is in B in virtue of the fact that s normalizes $\mathfrak{n}_0(E)$. But $\operatorname{ad}(s)$ permutes the roots of $\Delta(E)$ and hence $s^{-1}G(E)s = G(E)$. Since $n_1 \in G(E)$, $s^{-1}n_1 s \in G(E) \subset M(E)A \cdot N$, and we are done.

We now need the following lemma which will be proved later by different methods in Section 7.

LEMMA 3. *$M(E) \subset B_0(E) \cdot B$ and hence $M(E) \subset B(E)$.*

An immediate consequence is that $B(E) = M(E) \cdot A \cdot N$ which establishes the first decomposition for $B(E)$. Another consequence is that $B_0(E) \cdot B = B_0(E) \cdot Z = B(E)$. This proves the second decomposition formula for $B(E)$, and completes the proof of the theorem.

We remark in conclusion that $G/B(E)$ contains a dense open cell $P(E)$ which we may take as the projection of P into $G/B(E)$. Alternately it is easy to see that $P(E)$ is the one-one image of $N^-(E)$ in $G/B(E)$ so that $P(E)$ is identified with $\mathfrak{n}_0^-(E)$ by the map $N \to \exp(N)B(E)$. Moreover A acts on $P(E)$ just as $\operatorname{ad}(A)$ acts on $\mathfrak{n}_0^-(E)$, and the complement of $P(E)$ is of quasi-invariant measure 0.

Section 4. We preserve our previous notations. The homogeneous space G/K is a symmetric Riemannian space of non-compact type. Furstenberg defines a compactification of G/K using the notion of a boundary of G. Let $m^+(G/B)$ be the space of probability measures on G/B, upon which G acts as a transformation group. Since K is transitive on G/B, there is a unique $\mu \in m^+(G/B)$ invariant under K. Then $g \to g \cdot \mu$ maps G into $m^+(G/B)$, and since μ is K-invariant this becomes an equivariant mapping ϕ of G/K into $m^+(G/B)$. Then he shows that ϕ is one-one and $\overline{\phi(G/K)}$ may be viewed as a compactification of G/K. Similarly $G/B(E)$ carries a unique probability measure μ invariant under K and one can define in the same way an equivariant mapping ϕ_E of G/K into $m^+(G/B(E))$, the set of probability measures on $G/B(E)$. In general, however, ϕ_E need not be one-one. The next theorem considers this situation. We are indebted to R. Hermann for suggesting this line of proof. We first need a definition which is clearly inspired by [8]. We shall say that a subset of the fundamental roots of $\mathfrak{g}$ is a component, if it is the set of fundamental roots belonging to the complexification of a simple ideal of $\mathfrak{g}_0$. Two fundamental roots in different components cannot restrict to the same element of F so we may speak unambiguously of the components of F.

THEOREM 4. *The map* ϕ_E *of* G/K *into* $m^+(G/B(E))$ *is injective if and only if* E *does not contain any component of* F.

Proof. Let μ be the K-invariant probability measure on $G/B(E)$. Then ϕ_E is injective if and only if $g\mu = \mu$ implies $g \in K$. Now we may write $g = kak'$ with $k, k' \in K$ and $a \in \exp(\bar{\Lambda})$. Then if $kak' \cdot \mu = \mu$, we find that $a \cdot \mu = \mu$. Thus ϕ_E is injective if and only if $a \cdot \mu = \mu$ implies $a = 1$. Now a acts on $P(E) \cong \mathfrak{n}_0^-(E)$ linearly and as before we can introduce coordinates y_α such that $\operatorname{ad}(a)$ is diagonal with diagonal entries $\exp(\alpha(\log a))$, $\alpha \in \Sigma^-(E)$. But now if $S_\epsilon = \{p : y_\alpha(p) < \epsilon\}$, it is clear since $\alpha(\log a) \leqq 0$, that $a^{-1}S_\epsilon - S_\epsilon$ contains a non-void open set unless $\alpha(\log a) = 0$ for all $\alpha \in \Sigma^-(E)$. But μ vanishes on no non-void open set in $P(E)$ so that $\mu = a \cdot \mu$ is impossible unless $\alpha(\log a) = 0$ for all $\alpha \in \Sigma^-(E)$. Conversely if $\alpha(\log a) = 0$ for all $\alpha \in \Sigma^-(E)$, a acts as the identity transformation on $P(E)$ and hence on $G/B(E)$, and so $a \cdot \mu = \mu$.

Thus ϕ_E is not injective if and only if there is non-zero H in $\bar{\Lambda}$ such that $\alpha(H) = 0$ for all $\alpha \in \Sigma^-(E)$. Let F_i be the components of F, $i = 1, \cdots, n$. Then $\mathfrak{h}^- = \Sigma + \mathfrak{h}_i^-$ where $\mathfrak{h}_i^-$ is the annihilator of the F_j for $j \neq i$. Then $\Lambda = \prod_{i=1}^{n} \bar{\Lambda}_i$ where $\bar{\Lambda}_i$ is the closed positive Weyl chamber in $\mathfrak{h}_i^-$. If E contains a component of F, say F_1, let $0 \neq H \in \bar{\Lambda}_1$. Then $\alpha \in \Sigma^-(E)$, $\pi(\alpha) = \sum n_\gamma \gamma$, $\gamma \in F_j$, $j \neq 1$. Thus $\alpha(H) = 0$, and ϕ_E is not injective.

Conversely, let $0 \neq H \in \bar{\Lambda}$ be such that $\alpha(H) = 0$ for all $\alpha \in \Sigma^-(E)$. Then $H = \sum H_i$ with $H_i \in \Lambda_i$, and we may assume that $H_1 \neq 0$. Now suppose that we could find $\gamma \in F_1 - E$. Let α_1 be a fundamental root of $\mathfrak{g}$ such that $\pi(\alpha_1) = \gamma$. Let β be the highest root of a component of $\mathfrak{g}$ which F_1 defines. Then $\pi(\beta) = \sum n_\gamma \gamma$ $(\gamma \in F_1)$ with all $n_\gamma \geqq 1$. Since $\gamma(H_1) > 0$ for some $\gamma \in F_1$, $\beta(H_1) = \beta(H) > 0$. But clearly $-\beta \in \Sigma^-(E)$ and we have reached a contradiction. Therefore E must contain F_1, a component of F. Theorem 4 is now established.

If E satisfies the conditions of the previous theorem we shall say that E is faithful and that $G/B(E)$ is a faithful boundary.

If $H \in \Lambda$, a replica of the argument of Theorem 1 shows at $a_t \cdot \mu$ tends to a point mass at the origin in $G/B(E)$ where μ is the K invariant measure on $G/B(E)$. Now $\mathrm{ad}(k)a_tK$ for $H \in \Lambda$ and $k \in K$ are exactly the "regular" geodesics through the origin in G/K. Then $\phi_E(\mathrm{ad}(k)a_tK)$ converges to a point mass at $k \cdot B(E)$ in $G/B(E)$.

LEMMA 4. *If E is faithful, then the normal subgroup L of G acting trivially on $G/B(E)$ is the sum of the compact components of G, and the center of G.*

Proof. If $g \in L$, then g acts trivially on $m^+(G/B(E))$. Since ϕ_E is injective and equivariant, g acts trivially on G/K. Conversely if g acts trivially on G/K, let $\exp(tP)K$, $P \in \mathfrak{p}_0$ be a regular geodesic such that $\phi_E(\exp(tP)K)$ converges to a point mass at $b \in G/B(E)$. Since g fixes $\exp(tP)K$, g fixes the point mass at b in $G/B(E)$, and consequently G acts trivially on $G/B(E)$. Thus L is the subgroup of G which acts trivially on G/K and thus is the sum of the compact components of G and the center of G.

Section 5. We shall now examine the fine structure of $\overline{\phi_E(G/K)}$. We shall assume that E is faithful since it will become clear that the general case can be reduced to this one. If D is any subset of F, we denote by $O(E)^D$ the orbit of $B(D)$ through the origin in $G/B(E)$. Then $O(E)^D$ is given as $B(D)/B(D) \cap B(E)$. It is clear from Section 3 that $B(D) \cap B(E) = B(D \cap E)$. Furthermore, we know that $B(D) = G(D)Z \cdot L(D)$ and that $Z \cdot L(D) \subset Z \cdot L(D \cap E) \subset B(D \cap E)$. Thus $G(D)$ is transitive on $O(E)^D$ and

$$O(E)^D \cong G(D)/B(D \cap E) \cap G(D).$$

The subgroup in the denomenator is just the normalizer of $\mathfrak{n}_0(D \cap E)$ in $G(D)$. Since $\mathfrak{n}_0(D \cap E) = (\mathfrak{n}_0(D \cap E) \cap \mathfrak{g}_0(D)) + \mathfrak{n}_0(D)$, and since $G(D)$ normalizes $\mathfrak{n}_0(D)$, $B(D \cap E) \cap G(D)$ is the normalizer of $\mathfrak{n}_0(D \cap E) \cap \mathfrak{g}_0(D)$ in $G(D)$.

If we consider $G(D)$ as a semi-simple group, $\mathfrak{h}^-(D)$ is a maximal abelian subspace of $\mathfrak{g}_0(D) \ominus \mathfrak{k}_0(D)$ and D restricted to $\mathfrak{h}^-(D)$ forms a restricted fundamental system of $\mathfrak{g}_0(D)$, [8]. Moreover the positive nilpotent subalgebra of $\mathfrak{g}_0(D)$ corresponding to $E \cap D$, which we denote by $\mathfrak{n}_0(E \cap D)^D$, is just $\mathfrak{n}_0(E \cap D) \cap \mathfrak{g}_0(D)$. We have therefore proved the following lemma.

LEMMA 5. *D is a restricted fundamental system of $\mathfrak{g}_0(D)$ and $O(E)^D$ is the boundary of $G(D)$ corresponding to the subset $E \cap D$ of D.*

It follows that $O(E)^D$ contains a cell which we denote by $P(E)^D$. This cell may be identified with $\mathfrak{n}_0^-(E \cap D)^D$ by the map $n \to \exp(n) \cdot B(E)$.

It may happen that two different subsets of F yield the same orbit by this construction. We shall call a set D E-connected if no component of D (where D is viewed as a restricted fundamental system of $G(D)$) is contained in $E \cap D$. Note that this says that $O(E)^D$ is a faithful boundary of $G(D)$. We first need some elementary facts about E-connected sets. Since the proofs are straightforward and tedious we shall omit them.

If D' is any set of F, let D'' be the union of the components of D' which are contained in $F \cap D'$. Let $D = D' - D''$.

LEMMA 6. *D is the unique maximal E-connected subset of D'.*

LEMMA 7. *Let D be E-connected; then*

$$\begin{aligned} D'' &= \{\gamma : \gamma \in E \text{ and } \gamma \text{ is a component of } D \cup \{\gamma\}\} \\ &= \{\gamma : \gamma \in E \text{ and } D \cup \{\gamma\} \text{ is not } E\text{-connected}\}. \end{aligned}$$

Further $D' = D'' \cup D$ is the unique maximal subset of F which contains D as maximal E-connected set.

We now proceed to the principal results of this section.

THEOREM 5. (1) *If $D' \subset F$, and D is the maximal E-connected subset of D', then $O(E)^D = O(E)^{D'}$.*

(2) *If D and C are E-connected, and $O(E)^D = O(E)^C$, then $D = C$.*

(3) *If D is E-connected, then the largest subgroup of G mapping $O(E)^D$ onto itself is $B(D')$ where D' is defined in Lemma 7.*

(4) *If $g \in G$ maps $O(E)^D$ onto $O(E)^C$, then $O(E)^D = O(E)^C$.*

Proof. (1) D is a union of components of D' and so defines a union of components of a fundamental system $\alpha_1, \cdots, \alpha_m$ of $G(D')$; namely those α_i such that there is a root α of $G(D')$ with $\alpha = \sum n_j \alpha_j$ and $n_i > 0$ and $n_k > 0$ for some k such that $\pi(\alpha_k) \in D$. Let $\mathfrak{r}_0$ be the semi-simple ideal of $\mathfrak{g}_0(D')$ which is the sum of the other components. Then it is clear that $\mathfrak{r}_0$ and $\mathfrak{g}_0(D)$ span $\mathfrak{g}_0(D')$ and that $\mathfrak{g}_0(D) \cap \mathfrak{r}_0$ is the sum of the compact components of $\mathfrak{g}_0(D)$.

We claim that $\mathfrak{r}_0 \subset \mathfrak{b}_0(E \cap D')$. It suffices to show that $\mathfrak{r} \subset \mathfrak{b}(E \cap D')$ where $\mathfrak{r}$ is the complexification of $\mathfrak{r}_0$. But $\mathfrak{r} = (\mathfrak{r} \cap \mathfrak{h}) + \sum \boldsymbol{C} \cdot X_\alpha$ $(\alpha \in S)$ where S is some set of roots. We observe that $\mathfrak{h} \subset \mathfrak{b}(E \cap D')$, and if $\alpha \in S$, then $\pi(\alpha) = \sum n_\gamma \gamma, \gamma \in D'' \subset E \cap D'$. Thus $\alpha \in \Delta(E \cap D)$ and $X_\alpha \in \mathfrak{b}(E \cap D')$.

Let R be the analytic group corresponding to $\mathfrak{r}_0$. Then it is clear that $G(D') = R \cdot G(D) = G(D)R$. But now $O(E)^{D'} = G(D')/B(E \cap D) \cap G(D')$

and since $R \subset B(E \cap D')$, $O(E)^{D'} = G(D)/B(E \cap D) \cap G(D) = O(E)^D$. This argument also shows that R acts trivially on $O(E)^D$.

(2) The natural map of $\mathfrak{n}_0^-(E)$ into $G/B(E)$ is injective and $\mathfrak{n}_0^-(E \cap D)^D$ is the subset mapping into $O(E)^D$. It follows that if $O(E)^D = O(E)^C$, then $\mathfrak{n}^-(E \cap D)^D = \mathfrak{n}^-(E \cap C)^C$. But $\mathfrak{n}^-(E \cap D)^D = \sum \mathbf{C} \cdot X_\alpha$ over those roots α such that $\pi(\alpha) = -\sum n_\gamma \gamma$ $(\gamma \in D)$ with at least one $n_\gamma > 0$ for $\gamma \in D - E$.

Thus if D and C are E-connected and $O(E)^D = O(E)^C$, it is clear that $D - E = C - E$. We show now that $D \cap E = C \cap E$. If $\beta \in D \cap E$, then we can find, since D is E-connected, a root $\alpha \in \Delta(D)$ such that $\pi(\alpha) = -\sum n_\gamma \gamma$ with $n_\beta > 0$ and $n_\gamma > 0$ for some $\gamma \in D - E$. Thus $X_\alpha \in \mathfrak{n}^-(D \cap E)^D = \mathfrak{n}^-(C \cap E)^C$. Since $n_\beta > 0$, $\beta \in C \cap E$ and so $D \cap E \subset C \cap E$. The reverse inclusion follows by symmetry, and we are done.

(3) The subgroup of G which maps $O(E)^D$ into itself must contain $B(D)$ and hence is of the form $B(C')$ for $C' \supset D$. It is clear that $O(E)^D = O(E)^{C'}$. We assume that D is E-connected and let C be the maximal E-connected subset of C'. Then $O(E)^D = O(E)^C$ and so $D = C$ by part (2). Hence in the notation of Lemma 7, $C' \subset D'$ since D is the maximal E-connected of C'. But by part (1) $O(E)^{D'} = O(E)^D$ and so $B(D') \supset B(C')$ and hence $D' \supset C'$. Thus $C' = D'$ as desired.

(4) Suppose that $g \in G$ maps $O(E)^D$ onto $O(E)^C$ where $B(D)$ and $B(C)$ are subgroups of G leaving invariant these orbits. We may assume that g maps the origin into the origin; that is, $g \in B(E)$. Also if $b, b' \in B$, bgb' will do the same, and by Bruhat's lemma we make take $g = s$, an element of $\hat{M}$. Now it is clear that $sB(D)s^{-1} = B(C)$ and since $s \in K$, $sM(D)s^{-1} = M(C)$ using the first decomposition of Theorem 3. It follows at once that s maps the $M(D)$ invariant measure μ_D on $O(E)^D$ onto the $M(C)$ invariant measure μ_C on $O(E)^C$.

We shall now make use of an argument due to Satake, [8], p. 98. Let $H \in \mathfrak{h}^-$ and $g_t = \exp(tH)K$ be a geodesic in G/K. We shall see later that $\alpha(H) = 0$ for $\alpha \in \Delta(D)$ and $\alpha(H) < 0$ for $\alpha \in \Sigma^-(D)$ are sufficient conditions that $\phi_B(g_t)$ converge to μ_D (The $M(D)$ invariant measure on $O(E)^D$). Let H and H' be in the closed positive Weyl chamber $\bar{\Lambda}$ of $\mathfrak{h}^-$ and satisfy these conditions relative to D and C. Since $\operatorname{ad}(s)\mathfrak{b}(D) = \mathfrak{b}(C)$, and $\operatorname{ad}(s)$ permutes the roots of $\mathfrak{g}$, it follows that $\operatorname{ad}(s)\Sigma^-(D) = \Sigma^-(C)$ and $\operatorname{ad}(s)\Delta(D) = \Delta(C)$. But now it is clear that $H(\lambda) = (1 - \lambda)H + \lambda s^{-1}H'$ satisfies the conditions above for D for $\lambda : 0 \leqq \lambda \leqq 1$. Hence $\phi_B(g_t(\lambda))$ converges to μ_D where $g_t(\lambda) = \exp(tH(\lambda))K$. Now $H(0) = H \subset \bar{\Lambda}$ and $H(1) = s^{-1}H' \in s^{-1}\bar{\Lambda}$. Thus we can find a finite sequence $s_0 = e, s_1, \cdots, s_k = s^{-1}$ in the Weyl group of $\mathfrak{h}^-$

such that for any λ, $H(\lambda) \in s_i\Lambda$ for some i, and for each i, there is a λ_i such that $H(\lambda_i) \in s_i\bar{\Lambda} \cap s_{i+1}\bar{\Lambda}$. But then $s_i s_{i+1}^{-1}$ must fix $H(\lambda_i)$ and hence $g_i(\lambda_i)$. Hence $s_i s_{i+1}^{-1}(\mu_D) = \mu_D$. It follows then that $s\mu_D = \mu_D = \mu_C$, and $O(E)^D = O(E)^C$. Theorem 5 is now completely proved.

Let D be an E-connected set and let D' be the maximal set such that $O(E)^D = O(E)^{D'}$. Now $G(D)/K(D)$ is a symmetric space of non-compact type. We seek a homomorphism $w(E,D)$ of $B(D')$ into the group of isometries of $G(D)/K(D)$. This has been defined by Satake, [8], p. 97. We shall reproduce this here. We observe that Z normalizes $G(D)$ and $K(D)$ and $Z \cap G(D) \subset K(D)$. Further Z preserves the Killing form on $\mathfrak{g}_0(D)$. Thus $Z \cdot G(D)$ will act as a group of isometries of $G(D)/K(D)$. If R is the subgroup of $G(D')$ defined in part (1) of Theorem 5, $G(D) \cdot Z \cap (R \cdot L(D'))$ is a normal subgroup of $G(D) \cdot Z$ contained in $K(D)$ and hence acts trivially on $G(D)/K(D)$. We observe that

$$B(D')/R \cdot L(D') \cong G(D) \cdot Z / G(D) \cdot Z \cap (R \cdot L(D')) = J.$$

Now define $w(E,D) = 1$ on $R \cdot L(D')$ and $w(E,D)$ is determined naturally by the established action of J as a group isometries on $G(D)/K(D)$.

We know that $O(E)^D$ is a faithful boundary of $G(D)$ since D is E-connected. Let $\phi^D_{D \cap E}$ denote the embedding map of $G(D)/K(D)$ into measures on $O(E)^D$. We preserve the previous notation.

THEOREM 6. *$B(D')$ is the largest subgroup of G which leaves the range of $\phi^D_{D \cap E}$ invariant. Moreover*

$$b \cdot \phi^D_{D \cap E}\ (s) = \phi^D_{D \cap E}\ (w(E,D)(b)s)$$

for $b \in B(D')$, $s \in G(D)/K(D)$.

Proof. First if g leaves the range of $\phi^E_{D \cap E}$ invariant, then $g \cdot \mu_D = \phi^C_{D \cap E}\ (s)$ for some s, where μ_D is the $M(D)$ invariant measure on $O(E)^D$. But μ_D and $\phi^D_{D \cap E}\ (s)$ both have $O(E)^D$ as support. Thus g maps $O(E)^D$ onto itself, and by Theorem 5, $g \in B(D')$.

It suffices now to prove the formula stated in the theorem. If $b \in B(D')$, we can write $b = g \cdot z \cdot l$ where $g \in G(D)$, $z \in Z$, $l \in R \cdot L(D')$. It clearly suffices to check the formula for g, z, and l separately. Now $R \cdot L(D')$ is a normal subgroup of $B(D)$ contained in $B(E \cap D)$ and so it acts trivially on $O(E)^D$. Since $w(E,D)(l) = 1$, the formula is immediate if $b = l$. If $b = g \in G(D)$ the formula holds by definition of $\phi^E_{D \cap E}$. Finally if $g = z$, let $s = t \cdot K(D)$ for $t \in G(D)$. Then $z \cdot \phi^E_{D \cap E}\ (s) = zt \cdot \mu_D = ztz^{-1}\mu_D$ since $z \in M(D)$. But now $w(E,D)(z)(s) = (ztz^{-1})K(D)$ and so

$$\phi^D_{D\cap E}\ (w(E,D)(z)(s)) = ztz^{-1}\mu_D = z\cdot\phi^D_{D\cap E}\ (s)$$

as desired.

Section 6. We continue with the same notation. Our main theorem gives the orbit structure of Furstenberg's compactifications.

THEOREM 7. *The closure* $\overline{\phi_E(G/K)}$ *of the range of* ϕ_E *consists of the union of the* G *orbits of the finite set of measures* μ_D *as* D *runs through the* E*-connected sets. Further these orbits are disjoint from each other.*

Proof. Let μ be the K-invariant probability measure on $G/B(E)$. Then $\phi_E(K) = \mu$ and $\overline{\phi_E(G/K)}$ is the closure of the set of measures $g\cdot\mu$, $g \in G$. Let $g_n \in G$ such that $g_n\mu \rightarrow \mu''$. Now $g_n = k_n a_n k'_n$ with $k_n, k'_n \in K$, $a_n \in \exp(\bar{\Lambda})$. Then $g_n\mu = k_n a_n\mu$. We can choose a subsequence $\{n_i\}$ such $k_{n_i} \rightarrow k$ and $a_{n_i}\mu$ has a limit. Then $k_{n_i}a_{n_i}\mu \rightarrow \mu''$. We now need the following lemma.

LEMMA 8. *Let a sequence* μ_n *of measures on* $G/B(E)$ *converge to* μ, *and let* $g_n \rightarrow g$ *in* G. *Then* $g_n\cdot\mu_n \rightarrow g\cdot\mu$.

Proof. Let f be a continuous function on $G/B(E)$. Then it is easy to see that $f(g_n x) \rightarrow f(x)$ uniformly for $x \in G/B(E)$. Let $|\mu_n| \leqq M$ (M is finite by the uniform boundedness theorem). Then

$$\begin{aligned} &|\textstyle\int f\,d(g\cdot\mu) - \int f\,d(g_n\cdot\mu_n)| \\ &\qquad \leqq |\textstyle\int f(gx)\,d\mu(x) - \int f(gx)\,d\mu_n(x)| + \int |f(gx) - f(g_n x)|\,d\mu_n(x). \end{aligned}$$

The first term tends to zero, and the second term is majorized by $\sup|f(gx) - f(g_n x)|\cdot M$ which also tends to zero. Therefore the lemma is established.

We continue with the proof of the theorem. By the lemma,

$$\mu'' = k\cdot\lim(a_{n_i}\cdot\mu) = k\mu'.$$

We only need show that $\mu' \in G\cdot\mu_D$ for some D.

If $\mu' = \lim\exp(H_n)\mu$, we can choose a subsequence of H_n (which we denote by the same symbols) such that $\alpha(H_n)$ either tend to $+\infty$ or to a finite non-negative limit for any $\alpha \in \Sigma^+$. Let D be that subset of F such that $\gamma(H_n)$ tends to a finite limit. Now we can find a sequence H''_n in $\bar{\Lambda}$ such that $\gamma(H_n) = \gamma(H''_n)$ for $\gamma \in D$ and $\gamma(H''_n) = 0$ for $\gamma \notin D$. It follows then that H''_n tends to a finite limit H'' in Λ. Let $H'_n = H_n - H''_n$ so that $\gamma(H'_n) = 0$ for $\gamma \in D$ and $\gamma(H'_n) = \gamma(H_n) \geqq 0$ for $\gamma \notin D$. Then $H'_n \in \bar{\Lambda}$, and $\exp(H_n) = \exp(H''_n)\exp(H'_n)$. We may assume, if necessary, that $\exp(H'_n)\mu$ tends to a limit ν. Then by Lemma 8,

$$\mu' = \lim \exp(H_n)\mu = \exp(H'')\nu.$$

Therefore it suffices to prove the theorem for ν. We change our notation for convenience so that $H'_n = H_n$. We first observe that $H_n \in \mathfrak{h}'^-(D)$. If $k \in M(D)$ (the centralizer of $\mathfrak{h}'^-(D)$), then

$$k \cdot \nu = k \cdot \lim \exp(H_n)\mu = \lim \exp(H_n)k\mu = \nu.$$

Therefore ν is a probability measure on $G/B(E)$ invariant under $M(D)$.

Recall that $P(E)^D = \exp(\mathfrak{n}_0^-(E \cap D)^D) \cdot B(E)$ is a Euclidean cell in $O(E)^D$. We introduce coordinates $y'_\alpha, \alpha \in \Sigma^-(E)$ on $\mathfrak{n}_0^-(E) \cong P(E)$ such that $\mathrm{ad}(a)$ becomes diagonal with eigenvalues $\exp(\alpha(\log a))$ and such that $y'_\alpha(\alpha \in \Sigma^-(E \cap D)^D)$ are coordinates on $\mathfrak{n}_0^-(E \cap D)^D$. Let y_α denote the corresponding coordinates on $P(E)$. Now if $\alpha \in \Sigma^-(E \cap D)^D$, $\pi(\alpha) = \sum n_\gamma \gamma$ $(\gamma \in D)$ and $\alpha(H_n) = 0$. Thus $a_n = \exp(H_n)$ acts trivially on $P(E)^D \cong \mathfrak{n}_0^-(E \cap D)^D$. If $\alpha \notin \Sigma^-(E \cap D)^D$. Then $\pi(\alpha) = \sum n_\gamma \gamma$, where $n_\gamma \leqq 0$ and at least one $n_\beta < 0$ for $\beta \notin D$. Then

$$\exp(\alpha(H_n)) \leqq \exp(-\beta(H_n)) \text{ and } \beta(H_n) \to +\infty.$$

Therefore we can find a sequence γ_n tending to $+\infty$ such that $\exp(-\alpha(H_n)) \geqq r_n$ for any $\alpha \in \Sigma^-(E) - \Sigma^-(E \cap D)^D$.

Let R be a closed subset of $P(E)^D$ and let R' and R'' be two neighborhoods of R such that $R \subset R' \subset \bar{R}' \subset R''$. Now let $S(\epsilon, R') = \{x + z \colon x \in R'$ and $y_\alpha(z) = 0$ if $\alpha \in \Sigma^-(E \cap D)^D$ and $|y_\alpha(z)| < \epsilon.\}$. Now it is clear that we can find a continuous function on $P(E)$ such that $0 \leqq f \leqq 1$ with $f = 1$ on $S(\epsilon, R')$ and such that the support of f is contained in $S(2\epsilon, R'')$. Then

$$\int f\, d\nu = \lim \int f(a_n \cdot x)\, d\mu(x).$$

Since $\mathrm{ad}(a_n) = 1$ on $\mathfrak{n}_0^-(E \cap D)^D$ and all eigenvalues of $\mathrm{ad}(a_n^{-1})$ for $\alpha \notin \Sigma^-(E \cap D)^D$ exceed r_n, $f(a_n \cdot x) = 1$ on $S(r_n\epsilon, R')$ and $f(a_n \cdot x)$ is supported on $S(\infty, R'')$. Thus we find that

$$\mu(S(r_n\epsilon, R')) \leqq \int f(a_n x)\, d\mu(x) \leqq \mu(S(\infty, R''))$$

and thus

$$\mu(S(\infty, R')) \leqq \int f\, d\nu \leqq \mu(S(\infty, R''))$$

since $\mu(S(r_n\epsilon, R')) \uparrow \mu(S(\infty, R'))$. Finally as $R'' \downarrow R$

$$\mu(S(\infty, R'')) \downarrow \mu(S(\infty, R))$$

since μ is a finite measure. But as $R'' \downarrow R$ and $\epsilon \to 0$,

$$\int f\, d\nu \to \nu(R).$$

Thus we find $\nu(R) = \mu(S(\infty,R))$.

If $R \uparrow P(E)^D$, then $\mu(S(\infty,R)) \uparrow \mu(P(E)) = 1$ and so $\nu(P(E)^D) = 1$. Thus since ν is also a probability measure, ν must live on $P(E)^D$. Therefore the support of ν is contained in the closure $O(E)^D$ of $P(E)^D$. But $M(D)$ is transitive on $O(E)^D$ and ν is $M(D)$ invariant. It follows immediately that $\nu = \mu_D$. To be more precise, let C be the maximal E-connected subset of D, then $O(E)^D = O(E)^C$ and $\nu = \mu_D = \mu_C$.

We now must show that each $\mu_D \in \overline{\phi_B(G/K)}$. In fact let $H \in \bar{\mathrm{A}}$ such that $\alpha(H) = 0$ for $\alpha \in \Delta(D)$ and $\alpha(H) > 0$ for $\alpha \in \Sigma^+(D)$. Then the above argument shows that if $a_t = \exp(tH)$, then every subsequence of $a_t \cdot \mu$ which converges at all, must converge to μ_D. Thus $a_t \cdot \mu \rightarrow \mu_D$. In fact we can drop the condition that $H \in \bar{\mathrm{A}}$. This then clears up the point which was left open in the proof of part (4) of Theorem 5.

We further claim that the orbits $G \cdot \mu_D$ are disjoint. For if $g \cdot \mu_D = h \cdot \mu_C$ where D and C are E-connected, it is clear that hg^{-1} maps the support $O(E)^D$ of μ_D onto the support $O(E)^C$ of μ_C. By Theorem 5 (3), $D = C$. Theorem 7 is completely proved.

Finally we clear up Lemma 3 used in the proof of Theorem 3. We take $E = \phi$ and H an element of $\mathfrak{h}'^{-}(D)$ such that $\gamma(H) > 0$ for $\gamma \notin D$. If $a_t = \exp(tH)$, then our argument shows that $a_t \cdot \mu$ converges to a measure ν which lives on $\exp(\mathfrak{n}_0'^{-}(D)) \cdot B$ ($\mathfrak{n}_0'^{-}(D)$ is our notation for $\mathfrak{n}_0'^{-}(\phi)^D$.) and is $M(D)$ invariant. Moreover ν vanishes on no non-void open subset of this set. But it is clear that $\exp(\mathfrak{n}'_0(D)) \cdot B$ is contained in the orbit of $B_0(D)$ through the origin in G/B. But this orbit is clearly closed, so that ν is supported on $B_0(D) \cdot B$ in G/B, and the support of ν contains $\exp(\mathfrak{n}'_0{}^{-}(D)) \cdot B$. Since ν is $M(D)$ invariant it follows that $M(D) \subset B_0(D) \cdot B$ as desired.

Section 7. In order to show the relation of these compactifications with those of Satake [8], we will need two more lemmas of a topological nature.

LEMMA 9. *If E is faithful, ϕ_B is a homeomorphism of G/K into $m^+(G/B(E))$.*

Proof. If $g_n \in G$, and $g_n \cdot \mu$ converges to $g \cdot \mu$ (μ is the K-invariant measure on $G/B(E)$), we shall show that $g_n K \rightarrow gK$. If the set of elements gK, $g_n K$ remains in a compact set S of G/K, then ϕ_B is a homeomorphism on S, and it will follow that $g_n K \rightarrow gK$. Thus we write $g_n = k_n a_n k'_n$, and it will suffice to show that $\log(a_n) = H_n$ remains bounded in $\bar{\mathrm{A}}$. If not, we can find a subsequence which we may denote by H_n and a proper subset D of F such that $\gamma(H_n) \rightarrow +\infty$ for $\gamma \notin D$ and $\gamma(H_n)$ tends to a finite limit for $\gamma \in D$. We may also assume that $k_n \to k \in K$. We know that $a_n \cdot \mu$ has a limit ν which is supported on $O(E)^D$. Since E is faithful, $O(E)^D$ is a proper submanifold of $G/B(E)$. By Lemma 8, $g \cdot \mu = \lim g_n \cdot \mu = k\nu$ is supported on $k \cdot O(E)^D$ which is a proper submanifold. But $g \cdot \mu$ has support equal to all of $G/B(E)$ since it is quasi-invariant. This contradiction establishes the lemma..

Now let ψ_E denote the map of $\bar{A}$ into $m^+(G/B(E))$ given by $\psi_E(H) = \phi_E(\exp(H))$. Similarly we denote by $\psi^D_{E \cap D}$ the obvious map of $\bar{A}(D)$ (the positive chamber in $\mathfrak{h}^-(D)$) into $m^+(O(E)^D)$.

LEMMA 10. $\overline{\psi_E(\bar{A})} = \bigcup \psi^D_{E \cap D}(\bar{A}(D))$ *over all E-connected sets D. Furthermore* $\psi^D_{E \cap C}(H_n)$ *converges to* $\psi^C_{E \cap D}(H_0)$ *if and only if* $C \subset D$ *and* $\gamma(H_n) \to \gamma(H_0)$ *for* $\gamma \in C$, *and for every E-connected B such that* $C \subsetneq B \subset D$, *there is a* $\gamma \in B - C$ *with* $\gamma(H_n) \to +\infty$.

Proof. The first statement was essentially proved in Theorem 7. We claim further that $\psi^D_{E \cap C}(H_n) \to \psi^C_{E \cap C}(H_0)$ if and only if $C \subset D$ and $\gamma(H_n) \to \gamma(H_0)$ for $\gamma \in C$ and $\gamma(H_n) \to +\infty$ for $\gamma \notin C'$ where C' is the largest subset of D which contains C as maximal $E \cap D$-connected set. This also follows immediately from the proof of Theorem 7.

We only need show then that $\gamma(H_n) \to +\infty$, $\gamma \notin C'$ is equivalent to the final condition in the statement of the lemma. Let $\gamma(H_n) \to +\infty$ for $\gamma \notin C'$, and let $C \subsetneq B \subset D$ with B, $E \cap D$-connected. B cannot be contained in C' for otherwise $C = B$. Thus there exists a $\gamma \in B - C'$, and we have $\gamma(H_n) \to +\infty$. Conversely, let the condition be satisfied, and let $\gamma \notin C'$. Then $\{\gamma\} \cup C$ is $E \cap D$-connected. (If $\gamma \notin E \cap D$, this is clear, and if $\gamma \in E \cap D$, our assertion follows from Lemma 7.) Since $\beta(H_n)$ remains finite for $\beta \in C$, $\gamma(H_n)$ must tend to ∞, and we are done.

Satake [8], constructs a compactification of G/K for each subset I of F such that every component of F meets I. We denote this compactification by $(G/K)_I$. First we note that I satisfies Satake's condition if and only if $F - I$ is faithful in our sense.

THEOREM 8. *If E is faithful,* $\overline{\phi_E(G/K)}$ *is isomorphic to* $(G/K)_{F-E}$.

Proof. Satake gives a list of axioms 1°-4° on p. 100 which completely characterize $(G/K)_{F-E}$. We shall show that $\overline{\phi_E(G/K)}$ satisfies these axioms. It is clear that $\overline{\phi_E(G/K)}$ is a compact G-space and that G/K is homeomorphically embedded (Lemma 9) as a dense subspace.

For Satake's second condition, we note that if ρ is the highest weight of a representation and $(F - E) = \Delta_\rho^-$ is the corresponding subset of F (see bot p. 102, [8] for notation), then a subset D of F is ρ-open in Satake's definition if and only if D is E-connected in our sense. A perusal of the definitions shows that Satake's condition 2° and 3° are equivalent to Theorems 3, 5, 6 and 7. Finally 4° is just our Lemma 10.

Therefore the compactifications of G/K obtained by embedding into probability measures on a faithful boundary are in one-to-one correspondence in a simple way with Satake's compactifications.

UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

REFERENCES.

[1] F. Bruhat, "Sur les Représentations Induites des Groupes de Lie," *Bulletin de la Société Mathématique de France*, vol. 84 (1956), pp. 97-205.

[2] H. Furstenberg, "A Poisson formula for semi-simple Lie groups," *Annals of Mathematics*, vol. 77 (1963), pp. 335-386.

[3] Harish-Chandra, "Some applications of the universal enveloping algebra of a semi-simple Lie algebra," *Transactions of the American Mathematical Society*, vol. 70 (1951), pp. 28-29.

[4] R. Hermann, "A Poisson kernel for homogeneous spaces," *Proceedings of the American Mathematical Society*, vol. 12 (1961), pp. 892-899.

[5] ———, "Geometric aspects of potential theory in bounded symmetric domains. I," *Mathematische Annalen*, vol. 148 (1962), pp. 349-366; II, III *to appear.*

[6] K. Iwasawa, "On some types of topological groups," *Annals of Mathematics*, vol. 50 (1949), pp. 507-558.

[7] G. W. Mackey, "Induced representations of locally compact groups I," *Annals of Mathematics*, vol. 55 (1952), pp. 101-140.

[8] I. Satake, "On representations and compactifications of symmetric Riemannian spaces," *Annals of Mathematics*, vol. 71 (1960), pp. 77-110.

[9] Séminaire, "Sophus Lie," E. N. S., Paris, 1954.

[10] Séminaire Bourbaki, 1960-61, No. 206.

ON THE GRADED RING OF THETA-CONSTANTS.*

By JUN-ICHI IGUSA.

Introduction. Let g be a positive integer and let τ be a complex symmetric matrix of degree g with a positive-definite imaginary part. The set of such matrices forms a convex open subset of the $\frac{1}{2}\cdot g(g+1)$-dimensional complex vector space and (as a complex manifold) it is called the *Siegel upper-half plane* of degree (or "genus") g. A standard notation for this is $\mathfrak{S}_g$. Observe that $\mathfrak{S}_1$ is the upper-half of the ordinary complex plane. We know that the group $Sp(g,\boldsymbol{R})$ operates transitively on $\mathfrak{S}_g$ as

$$M\cdot\tau=\begin{pmatrix} a & b\\ c & d\end{pmatrix}\cdot\tau=(a\tau+b)(c\tau+d)^{-1}$$

and we get in this way (i.e. $Sp(g,\boldsymbol{R})/\pm 1_{2g}$ is) the full group of analytic automorphisms of $\mathfrak{S}_g$. Moreover discrete subgroups of $Sp(g,\boldsymbol{R})$ give rise to all possible groups of analytic automorphisms which operate properly discontinuously on $\mathfrak{S}_g$. Among discrete subgroups of $Sp(g,\boldsymbol{R})$, the importance of the *Siegel modular group* $Sp(g,\boldsymbol{Z})$ is unquestionable. We consider a discrete subgroup Γ of $Sp(g,\boldsymbol{R})$ which is "commensurable" with $Sp(g,\boldsymbol{Z})$ (in the sense that the intersection has finite indices in Γ and $Sp(g,\boldsymbol{Z})$). An analytic function ψ on $\mathfrak{S}_g$ which satisfies the functional equation

$$\psi(M\cdot\tau)=\det(c\tau+d)^k\psi(\tau)$$

for every M in Γ (plus a condition at infinity for $g=1$) is called a *Siegel modular form* or simply a modular form of weight k belonging to Γ. We assume that the weight is an integer and then we conclude (unless $\psi=0$) that it is non-negative. The set of such functions forms a complex vector space $A(\Gamma)_k$ and the graded ring

$$A(\Gamma)=\bigoplus_{0\leqq k<\infty} A(\Gamma)_k$$

will be called the *ring of modular forms* belonging to Γ. It is a normal integral domain and the theory of compactifications [1] shows that it is of

---

Received October 11, 1963.

* This work was partially supported by the National Science Foundation.

finite type over $A(\Gamma)_0 = \boldsymbol{C}$ and that the complex projective variety proj. $A(\Gamma)$ [cf. 4, footnote 2], which we simply call the *modular variety* belonging to Γ, contains the quotient variety $\Gamma\backslash\mathfrak{S}_g$ as a Zariski open set whose complement is of co-dimension (at least) two for $g \geqq 2$.

Now the problem we are interested in is the structure of the modular variety proj. $A(\Gamma)$ and in particular the one belonging to $Sp(g, \boldsymbol{Z})$. At present moment, there are three methods to investigate this problem. The first one is to study the variety of moduli of curves of genus g and then to use the information so obtained for the investigation of the modular variety. This method works only for $g = 1, 2, 3$, but we can use the classical theory of invariants of binary quartics, sextics or ternary quartics according to the cases. In fact, in this way we succeeded rather recently in determining the structure of proj. $A(Sp(2, \boldsymbol{Z}))$ [4, 5]. The second method is to study the variety of moduli of principally polarized abelian varieties. This method works for all g and we can use another legacy of the past, i. e. the theory of theta-functions. The third method is to work out a finite form of the Selberg trace formula and to apply this formula to the investigation of proj. $A(\Gamma)$. This, when worked out successfully, will be the most general method. The purpose of this paper is to establish what we think is a fundamental lemma in the second method. In the following we shall give its exact statement.

We shall write vectors in $\boldsymbol{Z}^{2g}$ by $\mathfrak{m}, \cdots$ and we use m', m'' to denote vectors in $\boldsymbol{Z}^g$ determined by the first and the last g coefficients of $\mathfrak{m}$. Now, if τ is a point of $\mathfrak{S}_g$ and z a point of $\boldsymbol{C}^g$, the following series

$$\theta_{\mathfrak{m}}(\tau, z) = \sum_{p \in \boldsymbol{Z}^g} e[\tfrac{1}{2} \cdot {}^t(p + m'/2)\tau(p + m'/2) \\ + {}^t(p + m'/2)(z + m''/2)]$$

converges absolutely and uniformly in every compact subset of $\mathfrak{S}_g \times \boldsymbol{C}^g$. Therefore, for a fixed $\mathfrak{m}$, it represents an analytic function of the two variables τ and z, which is called the *theta-function* of *characteristic* $\mathfrak{m}$. Actually, since we have

$$\theta_{\mathfrak{m}+2\mathfrak{n}}(\tau, z) = (-1)^{{}^t m' n''} \theta_{\mathfrak{m}}(\tau, z)$$

for any $\mathfrak{n}$ in $\boldsymbol{Z}^{2g}$, we sometimes normalize the characteristic by the condition that its coefficients are either zero or one. At any rate, if we put $z = 0$, we get the so-called *theta-constants* $\theta_{\mathfrak{m}}(\tau) = \theta_{\mathfrak{m}}(\tau, 0)$ and $\theta_{\mathfrak{m}}$, which we also call theta-constants, are analytic functions in $\mathfrak{S}_g$. On the other hand, for every positive integer n, we consider the subgroup $\Gamma_g(n, 2n)$ of $Sp(g, \boldsymbol{Z})$ of elements M satisfying

$$M \equiv 1_{2g} \bmod n, \qquad (a\,{}^t b)_0 \equiv (c\,{}^t d)_0 \equiv 0 \bmod 2n.$$

In general, if s is a square matrix, its diagonal coefficients arranged in a natural order form a vector which is denoted by $(s)_0$. The subgroup $\Gamma_g(1,2)$ appeared already in the analytic theory of quadratic forms. At any rate, we shall show that *the graded ring* $A(\Gamma_g(4,8))$ *is the normalization of the graded ring* $\boldsymbol{C}[\theta_{\mathrm{m}}\theta_{\mathrm{n}}]$ *generated over* $\boldsymbol{C}$ *by the products of theta-constants.* Namely the two integral domains have the same field of fractions and $A(\Gamma_g(4,8))$ is the "integral closure" of $\boldsymbol{C}[\theta_{\mathrm{m}}\theta_{\mathrm{n}}]$. If g is given, we can certainly determine the normalization of $\boldsymbol{C}[\theta_{\mathrm{m}}\theta_{\mathrm{n}}]$ explicitly and thus we will know the ring $A(\Gamma_g(4,8))$. Then, by a going-down process, which is *linear,* we will know the ring $A(Sp(g,\boldsymbol{Z}))$. Actually, in the two cases $g=1,2$, the ring $\boldsymbol{C}[\theta_{\mathrm{m}}\theta_{\mathrm{n}}]$ itself is normal, and in the case $g=2$, we can reproduce our structure theorem with further important information. This will be discussed in a separate paper. The comparison in the case $g=2$ shows that the second method is considerably simpler than the first method. It is free from the theory of invariants and we need not "blow up" anything (which was necessary to fill the gap between the two types of varieties of moduli). Therefore, it is expected that the calculation in the case $g=3$, for instance, is something of a manageable order.

1. **Congruence group $\Gamma_g(n,2n)$.** As we already did in the introduction, we shall express a typical element M of $Sp(g,\boldsymbol{R})$ in the form

$$M=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

by four $g\times g$ submatrices a, b, c, d. We recall that $Sp(g,\boldsymbol{R})$ is the group of $2g\times 2g$ real matrices M satisfying $ME{}^tM=E$ in which

$$E=\begin{pmatrix} 0 & 1_g \\ -1_g & 0 \end{pmatrix},$$

and $Sp(g,\boldsymbol{Z})$ is its discrete subgroup of elements with integer coefficients. Observe that the condition $ME{}^tM=E$ can be written in terms of a, b, c, d as

$$a{}^tb=b{}^ta, \qquad c{}^td=d{}^tc, \qquad a{}^td-b{}^tc=1_g.$$

If n is a positive integer, the *principal congruence group* of degree g and of level n is defined as a subgroup of $Sp(g,\boldsymbol{Z})$ by the condition $M\equiv 1_{2g} \bmod n$. We shall denote this group by $\Gamma_g(n)$. Observe that $\Gamma_g(n)$ is stable by $M\to {}^tM$ and that it is normal in $\Gamma_g(1)=Sp(g,\boldsymbol{Z})$. In general, a subgroup of $\Gamma_g(1)$ is called a congruence group (of degree g and) of level n if it contains $\Gamma_g(n)$. The group $\Gamma_g(n,2n)$ in the introduction is, therefore, a congruence group of level $2n$. We shall use the following formula

$$[\Gamma_g(1):\Gamma_g(n)]=n^{g(2g+1)}\prod_{p|n}\prod_{1\leqq k\leqq g}(1-p^{-2k}),$$

which we can find, for instance, in Koecher [6]. We observe that, if M is an element of $\Gamma_g(n)$ and n is even, we have $(a{}^tb)_0\equiv(b)_0$, $(c{}^td)_0\equiv(c)_0 \bmod 2n$. Therefore, in case n is even, we can define $\Gamma_g(n,2n)$ as a subgroup of $\Gamma_g(n)$ by the condition $(b)_0\equiv(c)_0\equiv 0 \bmod 2n$. In particular, it is stable by $M\rightarrow{}^tM$. We shall prove the following lemma:

LEMMA 1. *If n is even, then* (i) *$\Gamma_g(n,2n)$ is a normal subgroup of $\Gamma_g(1)$ and its index in $\Gamma_g(n)$ is 2^{2g},* (ii) *$\Gamma_g(n)/\Gamma_g(2n,4n)$ is an abelian group,* (iii) *$\Gamma_g(n,2n)/\Gamma_g(2n,4n)$ is a vector space over $\mathbf{Z}/2\mathbf{Z}$ of dimension $g(2g+1)$.*

Proof. If we express a typical element M of $\Gamma_g(n)$ in the form

$$1_{2g}+n\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$$

with $g\times g$ integer matrices α, β, γ, δ, the condition for this element to be in $\Gamma_g(n)$ is

$${}^t\beta=\beta+n(\beta{}^t\alpha-\alpha{}^t\beta),\qquad {}^t\gamma=\gamma+n(\gamma{}^t\delta-\delta{}^t\gamma)$$
$$\alpha+{}^t\delta+n(\alpha{}^t\delta-\beta{}^t\gamma)=0.$$

We associate to this element of $\Gamma_g(n)$, a vector over $\mathbf{Z}/2\mathbf{Z}$ whose coefficients are $(\beta)_0 \bmod 2$ and $(\gamma)_0 \bmod 2$. Then we get a homomorphism of $\Gamma_g(n)/\Gamma_g(2n)$ to $(\mathbf{Z}/2\mathbf{Z})^{2g}$. Since we can take $\alpha=\gamma=\delta=0$, $\beta=$ arbitrary (integer) symmetric matrix, this homomorphism is surjective. Since the kernel is precisely $\Gamma_g(n,2n)/\Gamma_g(2n)$, the index of $\Gamma_g(n,2n)$ in $\Gamma_g(n)$ is 2^{2g}. Now we transform M by an arbitrary element of $\Gamma_g(1)$ composed of a, b, c, d. Then we get

$$1_{2g}+n\begin{pmatrix}\alpha' & \beta'\\ \gamma' & \delta'\end{pmatrix}$$

with $\beta'=-(a\alpha+b\gamma){}^tb+(a\beta+b\delta){}^ta$, hence

$$(\beta')_0\equiv a(\beta)_0+b(\gamma)_0 \bmod 2.$$

Therefore $(\beta)_0\equiv(\gamma)_0\equiv 0 \bmod 2$ implies $(\beta')_0\equiv 0 \bmod 2$ and also $(\gamma')_0\equiv 0 \bmod 2$ (by a similar argument or by using transposed matrices). Hence $\Gamma_g(n,2n)$ is normal in $\Gamma_g(1)$, and this proves (i). We now consider two elements M_i of $\Gamma_g(n)$ determined by α_i, β_i, γ_i, δ_i for $i=1,2$. Then their commutator is of the form

$$1_{2g}+n\begin{pmatrix}\alpha_3 & \beta_3\\ \gamma_3 & \delta_3\end{pmatrix}$$

with

$$\beta_3 \equiv n(\alpha_1{}^t\beta_2 + \beta_2{}^t\alpha_1 - \alpha_2{}^t\beta_1 - \beta_1{}^t\alpha_2) \bmod n^2,$$

hence $(\beta_3)_0 \equiv 0 \bmod 4$ and also $(\gamma_3)_0 \equiv 0 \bmod 4$. Since $\Gamma_g(n)/\Gamma_g(2n)$ is an abelian group, therefore, the commutator group of $\Gamma_g(n)$ is contained in $\Gamma_g(2n, 4n)$. This proves (ii). As for (iii), we first observe that squares of elements of $\Gamma_g(n, 2n)$ are contained in $\Gamma_g(2n, 4n)$. Then we have only to recall that the indices of $\Gamma_g(2n)$ in $\Gamma_g(n)$ and of $\Gamma_g(4n)$ in $\Gamma_g(2n)$ are both equal to $2^{g(2g+1)}$, and apply (i). This completes the proof.

We note that, in this lemma, the assumption that n is even is essential. In fact, otherwise all three statements are false. Now we turn our attention to $\Gamma_g(2)$ and construct a set of generators.

If i, j are distinct positive integers at most equal to g, we replace the (i, j)-coefficient in 1_g by 2. We take this matrix as a and put $b = c = 0$, $d = {}^ta^{-1}$. The corresponding matrix A_{ij} is then an element of $\Gamma_g(2)$. If i is a positive integer at most equal to g, we replace the (i, i)-coefficient in 1_g by -1. We take this matrix as a and put $b = c = 0$, $d = {}^ta^{-1}$. The corresponding matrix A_{ii} is an element of $\Gamma_g(2)$. If i, j are integers satisfying $1 \leqq i < j \leqq g$, we replace the (i, j)- and (j, i)-coefficients in 0 by 2. We take this matrix as b and put $a = d = 1_g$, $c = 0$. The corresponding matrix B_{ij} is then an element of $\Gamma_g(2)$. If i is a positive integer at most equal to g, we replace the (i, i)-coefficient in 0 by 2. We take this matrix as b and put $a = d = 1_g$, $c = 0$. The corresponding matrix B_{ii} is an element of $\Gamma_g(2)$. Finally we put $C_{ij} = {}^tB_{ij}$ for $i \leqq j$. We shall prove the following theorem:

THEOREM 1. *The* $g(2g+1)$ *matrices* A_{ij} $(1 \leqq i, j \leqq g)$, B_{ij}, C_{ij} $(1 \leqq i \leqq j \leqq g)$ *form a set of generators of* $\Gamma_g(2)$.

Proof. Suppose that M is a given element of $\Gamma_g(2)$. Since the diagonal coefficients of M are odd while other coefficients of M are even, they are different even in absolute values. This being remarked, we multiply suitable powers of $A_{12}, \cdots, A_{1g}, B_{1g}, \cdots, B_{11}$ in this order from the right to M and make the coefficients $a_{12}, \cdots, a_{1g}, b_{11}, \cdots, b_{1g}$ less than abs. a_{11} in absolute values. In case they are not all zero, suppose that a_{1i} or possibly b_{1j} is the first coefficient in the sequence which is different from zero. Then multiplying a power of A_{i1} or C_{1j} from the right according to the case, we make a_{11} less than abs. a_{1i} or abs. b_{1j} in absolute value. Applying these two operations alternatively, we make $a_{12}, \cdots, a_{1g}, b_{11}, \cdots, b_{1g}$ all zero. Then we get $a_{11}d_{11} = 1$, hence by multiplying A_{11} if necessary, we get $a_{11} = d_{11} = 1$. This being done, we multiply suitable powers of $A_{21}, \cdots, A_{g1}, C_{11}, \cdots, C_{1g}$ in this order from the left and (without changing the first line) make $a_{21}, \cdots, a_{g1}$,

$c_{11}, \cdots, c_{g1}$ all zero. In this way M is expressed as a product of powers of the A_{ij}, B_{ij}, C_{ij} (more precisely of A_{1j}, A_{i1}, B_{1j}, C_{1j}) and the following matrix

$$M^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & a_1 & b'' & b_1 \\ 0 & c' & 1 & d' \\ 0 & c_1 & d'' & d_1 \end{bmatrix}.$$

Since M^* is in $\Gamma_g(1)$ (in fact in $\Gamma_g(2)$), we have

$$\begin{aligned} & b'' = 0, \ a_1{}^t b_1 = b_1{}^t a_1 \\ & c'{}^t d' = d'{}^t c', \ c'{}^t d_1 = d'{}^t c_1, \ c_1{}^t d_1 = d_1{}^t c_1 \\ & d'' = 0, \ a_1{}^t d' = b_1{}^t c', \ a_1{}^t d_1 - b_1{}^t c_1 = 1_{g-1}. \end{aligned}$$

Since ${}^tM^*$ is of the same form as M^* and is in $\Gamma_g(1)$, we have not only $b'' = d'' = 0$ but also $c' = d' = 0$. Moreover

$$M_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$$

is in $\Gamma_{g-1}(1)$, in fact in $\Gamma_{g-1}(2)$. We have only to observe, finally, that our set of matrices is such that we can apply an induction on the degree. This completes the proof.

We note that, since for any prime factor p of n, the factor group $\Gamma_g(n)/\Gamma_g(np)$ is commutative and indeed a vector space over $\boldsymbol{Z}/p\boldsymbol{Z}$ of dimension $g(2g+1)$, no $\Gamma_g(n)$ for $n > 1$ is generated by less than $g(2g+1)$ elements. In particular, none of the generators of $\Gamma_g(2)$ in Theorem 1 is redundant. At any rate, Theorem 1 implies that, if M is an arbitrary element of $\Gamma_g(2)$, it can be written in the form

$$M = \prod_{1 \leqq i,j \leqq g} A_{ij}{}^{p_{ij}} \prod_{1 \leqq i \leqq j \leqq g} B_{ij}{}^{q_{ij}} \prod_{1 \leqq i \leqq j \leqq g} C_{ij}{}^{r_{ij}} \cdot M'$$

with p_{ij}, q_{ij}, r_{ij} in $\boldsymbol{Z}$ and M' in the commutator group of $\Gamma_g(2)$, which are by no means unique. We shall denote the $g \times g$ integer matrix (p_{ij}) by p. Similarly we shall use q, r to denote the symmetric matrices defined by q_{ij}, r_{ij}. Using this notation, we shall prove the following statement:

SUPPLEMENT. *Suppose that an element*

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

of $\Gamma_g(2)$ is expressed as a product of powers of A_{ij}, B_{ij}, C_{ij} with exponents given by p, q, r and of an element of the commutator group of $\Gamma_g(2)$. Then

M is in $\Gamma_g(2,4)$, $\Gamma_g(4)$ or in $\Gamma_g(4,8)$ if and only if $(q)_0 \equiv (r)_0 \equiv 0 \bmod 2$, $p, q, r \equiv 0 \bmod 2$ *or* $(q)_0 \equiv (r)_0 \equiv 0 \bmod 4$, $p, q, r \equiv 0 \bmod 2$.

Proof. We observe that the said product of powers of A_{ij}, B_{ij}, C_{ij} is congruent to the following matrix

$$\begin{pmatrix} \begin{matrix} (-1)^{p_{11}} \cdots 2p_{1g} \\ \cdots\cdots \\ 2p_{g1} \cdots (-1)^{p_{gg}} \end{matrix} & 2q \\ 2r & \begin{matrix} (-1)^{p_{11}} \cdots 2p_{g1} \\ \cdots\cdots \\ 2p_{1g} \cdots (-1)^{p_{gg}} \end{matrix} \end{pmatrix}$$

modulo the greatest common divisor of the coefficients of $4p$, $4q$, $4r$. We recall also that M' is contained in $\Gamma_g(4,8)$ by Lemma 1, (ii). The supplement follows immediately from these two facts.

2. Transformation law of theta-functions. The definition of theta-functions was already recalled in the introduction. We shall review some elementary properties of the theta-functions with brief indications of proofs [cf. 7, 8]. First of all, as an analytic function of z, the theta-function $\theta_{\mathfrak{m}}(\tau, z)$ is characterized up to a constant factor by the functional equation

$$\theta_{\mathfrak{m}}(\tau, z + (\tau 1_g)\mathfrak{n}) = e(\mathfrak{m}, \mathfrak{n})\, e(-{}^t n' z - \tfrac{1}{2} \cdot {}^t n' \tau n')\, \theta_{\mathfrak{m}}(\tau, z)$$

satisfied for every $\mathfrak{n}$ in $\boldsymbol{Z}^{2g}$ where

$$e(\mathfrak{m}, \mathfrak{n}) = (-1)^{{}^t m' n'' - {}^t m'' n'}$$

and $e(\ )$ stands, as usual, for $\exp(2\pi i\ )$. In fact, the theta-function satisfies the functional equation. Conversely, if $\theta(z)$ satisfies the functional equation, we see that $e(-\tfrac{1}{2} \cdot {}^t m' z)\theta(z)$ is invariant under $z \to z + n''$. Expanding this periodic function into a Fourier series, we use the functional equation for $n'' = 0$, and we get $\theta(z) = \text{const.}\ \theta_{\mathfrak{m}}(\tau, z)$. We next observe that, as an analytic function of τ and z, the theta-function is characterized up to a constant factor by the functional equation and by the heat equation

$$\sum_{1 \leqq j, k \leqq g} \gamma_{jk} \partial^2 \theta_{\mathfrak{m}} / \partial z_j \partial z_k = 4\pi i \sum_{1 \leqq j \leqq k \leqq g} \gamma_{jk} \partial \theta_{\mathfrak{m}} / \partial \tau_{jk}$$

satisfied for any complex coefficients $\gamma_{jk} = \gamma_{kj}$. These being remarked, let λ be a complex non-singular matrix and μ a complex symmetric matrix both of degree g. Also let $\theta_{\mathfrak{n}}(\tau', z)$ be a theta-function of characteristic $\mathfrak{n}$. Then we have a relation of the form

$$\theta_{\mathfrak{n}}(\tau', z') = K \cdot e(\tfrac{1}{2} \cdot {}^t z \mu z) \theta_{\mathfrak{m}}(\tau, z) \quad (K = \text{const.})$$

as functions of $z' = \lambda z$ if and only if there exists an element M of $\Gamma_g(1)$ (composed of a, b, c, d) such that

$$\lambda = {}^t(c\tau + d)^{-1}, \quad \mu = (c\tau + d)^{-1} c$$
$$\tau' = M \cdot \tau, \quad \mathfrak{n} \equiv M \cdot \mathfrak{m} \mod 2,$$

where $M \cdot \tau$ was defined in the introduction and $M \cdot \mathfrak{m}$ has the following meaning

$$M \cdot \mathfrak{m} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \mathfrak{m} + \begin{pmatrix} (c\,{}^t d)_0 \\ (a\,{}^t b)_0 \end{pmatrix} .$$

In fact, under these conditions we can verify easily that $e(\tfrac{1}{2} \cdot {}^t z \mu z) \theta_{\mathfrak{m}}(\tau, z)$ satisfies the same functional equation as $\theta_{\mathfrak{n}}(\tau', z')$, hence they differ only by a constant factor. The proof of the necessity of the condition is more involved. We have to use the "primitivity" of the period group generated by the $2g$ column vectors of $(\tau 1_g)$ for $e(\tfrac{1}{2} \cdot {}^t z \mu z) \theta_{\mathfrak{m}}(\tau, z)$.

Concerning $\tau \rightarrow M \cdot \tau$ and $\mathfrak{m} \rightarrow M \cdot \mathfrak{m}$, we observe that they are defined for M in $Sp(g, \boldsymbol{R})$ and $\mathfrak{m}$ in $\boldsymbol{R}^{2g}$, and we have $1_{2g} \cdot \tau = \tau$, $1_{2g} \cdot \mathfrak{m} = \mathfrak{m}$. Moreover, if M_1, M_2 are in $Sp(g, \boldsymbol{R})$, we have $M_1 \cdot (M_2 \cdot \tau) = M_1 M_2 \cdot \tau$ and if they are in $Sp(g, \boldsymbol{Z})$, we have

$$M_1 \cdot (M_2 \cdot \mathfrak{m}) \equiv M_1 M_2 \cdot \mathfrak{m} \mod 2.$$

In particular $\Gamma_g(1)$ operates on the additive group of characteristics mod 2, which is $(\boldsymbol{Z}/2\boldsymbol{Z})^{2g}$, just as $Sp(g, \boldsymbol{R})$ operates on $\mathfrak{S}_g$. However, one is transitive and the other is not. In fact, if we put

$$e(\mathfrak{m}) = (-1)^{{}^t m' m''}$$

for any characteristic $\mathfrak{m}$, we have $e(M \cdot \mathfrak{m}) = e(\mathfrak{m})$ for every M in $\Gamma_g(1)$. We know, however, that $\Gamma_g(1)$ operates transitively on each set of characteristics mod 2 classified by $e(\mathfrak{m}) = \pm 1$. These characteristics are called even or odd because we have

$$\theta_{\mathfrak{m}}(\tau, -z) = e(\mathfrak{m}) \theta_{\mathfrak{m}}(\tau, z).$$

At any rate, the transitivity can be proved, for instance, by an induction on the degree. Now, if $\mathfrak{m}$ is odd, clearly the theta-constant $\theta_{\mathfrak{m}}(\tau) = \theta_{\mathfrak{m}}(\tau, 0)$ vanishes identically as a function of τ. On the other hand, if $\mathfrak{m}$ is even, it does not vanish identically. Since we have $\theta_{M \cdot \mathfrak{m}}(M \cdot \tau) = K \theta_{\mathfrak{m}}(\tau)$ and since K is different from zero, we have only to show that $\theta_0(\tau)$ does not vanish identically. But this is a consequence of

$$\lim_{\eta\to\infty}\theta_0(i\eta 1_g)=1.$$

These being remarked, we shall try to determine the constant K as explicitly as possible. At any rate K is a non-vanishing analytic function in $\mathfrak{S}_g$. Moreover, using the heat equation and the following relation

$$\delta\tau'={}^t(c\tau+d)^{-1}(\delta\tau)(c\tau+d)^{-1}$$

between differentials, we can derive the differential equation

$$\sum_{1\leqq j\leqq k\leqq g}\gamma_{jk}\partial\log K/\partial\tau_{jk}=\tfrac{1}{2}\cdot\sum_{1\leqq j,k\leqq g}\gamma_{jk}\mu_{jk}$$

whenever the characteristic $\mathfrak{m}$ is even. On the other hand, the formal identity $\delta(\log(\det X))=\operatorname{trace}(X^{-1}\delta X)$ applied to $X=c\tau+d$, shows that $\det(c\tau+d)^{\frac{1}{2}}$ satisfies the same differential equation as K. In this way, we get

$$K=\text{const.}\det(c\tau+d)^{\frac{1}{2}}.$$

Observe that the square root is not unique but it is a well-defined non-vanishing analytic function in $\mathfrak{S}_g$. The constant we have now does no longer depend on τ.

We shall separate the part of the constant which depends on $\mathfrak{m}$. We observe that we have

$$\theta_{\mathfrak{m}}(\tau,z+\tfrac{1}{2}\cdot(\tau 1_g)\mathfrak{n})=e(-1/8\cdot{}^tn'\tau n'\\ -\tfrac{1}{2}\cdot{}^tn'(z+\tfrac{1}{2}\cdot(m''+n'')))\theta_{\mathfrak{m}+\mathfrak{n}}(\tau,z)$$

for any $\mathfrak{n}$ in $\boldsymbol{Z}^{2g}$. Also, for a given M, if we choose the sign of the square root $\det(c\tau+d)^{\frac{1}{2}}$, we have

$$\theta_{M\cdot 0}(\tau',z')=\varkappa(M)\ \det(c\tau+d)^{\frac{1}{2}}e(\tfrac{1}{2}\cdot{}^tz\mu z)\theta_0(\tau,z)$$

in which $\varkappa(M)$ depends only on M. We replace z by $z+\frac{1}{2}\cdot(\tau 1_g)\mathfrak{m}$ and make a little calculation. Then we get the following transformation law of theta-functions

$$\theta_{M\cdot\mathfrak{m}}(M\cdot\tau,{}^t(c\tau+d)^{-1}z)=\varkappa(M)e(\phi_{\mathfrak{m}}(M))\det(c\tau+d)^{\frac{1}{2}}\\ \cdot e(\tfrac{1}{2}\cdot{}^tz(c\tau+d)^{-1}cz)\theta_{\mathfrak{m}}(\tau,z)$$

in particular

$$\theta_{M\cdot\mathfrak{m}}(M\cdot\tau)=\varkappa(M)e(\phi_{\mathfrak{m}}(M))\det(c\tau+d)^{\frac{1}{2}}\theta_{\mathfrak{m}}(\tau),$$

with

$$\phi_{\mathfrak{m}}(M)=-\tfrac{1}{8}\cdot({}^tm'{}^tbdm'+{}^tm''{}^tacm''-2{}^tm'{}^tbcm''\\ -2{}^t(a{}^tb)_0(dm'-cm'')).$$

The constant $\text{æ}(M)$ has been explicitly calculated when the left bottom submatrix c is non-singular [7, pp. 180-2]. Also, using some of Siegel's calculations [10], we can obtain an explicit (but complicated) finite form for $\text{æ}(M)$ in general. In the next section, we shall proceed differently and determine $\text{æ}(M)^2$ explicitly for every M in $\Gamma_g(2)$.

3. Theta-constants as modular forms. We know that $\Gamma_g(2)$ is generated by $g(2g+1)$ matrices A_{ij}, B_{ij}, C_{ij}. We shall calculate the constant æ for these generators of $\Gamma_g(2)$. In order to increase the readability, we shall arrange the material so that it is self-contained. We shall start proving the following classical lemma:

LEMMA 2. *If w is a point of the upper-half plane and if the square root $(-iw)^{\frac{1}{2}}$ is the one which is positive when w is pure imaginary, we have*

$$\sum_{p\in Z} e(\tfrac{1}{2}\cdot(-1/w)p^2) = (-iw)^{\frac{1}{2}} \sum_{p\in Z} e(\tfrac{1}{2}\cdot wp^2).$$

Proof. If we apply the transformation law of theta-functions to the case $g=1$, $M=E$ (in Section 1), we see that the two sides differ at most by a constant factor. By specializing w to i, we see that the constant is one. This proves the lemma.

In the formula just proved, we replace w by $w-2$ and the new w by $-1/w$. Then we get the following formula

$$\sum_{p\in Z} e(\tfrac{1}{2}\cdot w/(2w+1)\cdot p^2) = (2w+1)^{\frac{1}{2}} \sum_{p\in Z} e(\tfrac{1}{2}\cdot wp^2),$$

in which $(2w+1)^{\frac{1}{2}}$ is in the first quadrant. We shall obtain one more formula. We apply the transformation law of theta-functions to the case $g=2$, $M=E$. Then we get

$$\begin{aligned}\sum_{p_1,p_2\in Z} e(1/2(\tau_{11}\tau_{22}-(\tau_{12})^2)\cdot(-\tau_{22}(p_1)^2-\tau_{11}(p_2)^2+2\tau_{12}p_1p_2))\\ = \text{const.}((\tau_{12})^2-\tau_{11}\tau_{22})^{\frac{1}{2}} \sum_{p_1,p_2\in Z} e(\tfrac{1}{2}\cdot(\tau_{11}(p_1)^2\\ +\tau_{22}(p_2)^2+2\tau_{12}p_1p_2)),\end{aligned}$$

in which $((\tau_{12})^2-\tau_{11}\tau_{22})^{\frac{1}{2}}$ is positive when τ_{11}, τ_{22} are pure imaginary and τ_{12} is real. We put $\tau_{11}=\tau_{22}=w$, $\tau_{12}=0$ in this formula and compare that with the square of the one in Lemma 2. In this way, we see that the constant is one. We then replace τ_{11}, τ_{22} by w_1, w_2 and τ_{12} by -2, where w_1, w_2 are points of the upper-half plane. In the formula so obtained, we replace w_1, w_2 by $-1/w_1$, $-1/w_2$. Then, using Lemma 2 again, we get

$$\sum_{p_1,p_2\in \mathbf{Z}} e(1/2(1-4w_1w_2)\cdot(w_1(p_1)^2+w_2(p_2)^2-4w_1w_2p_1p_2)) \\ =(1-4w_1w_2)^{\frac{1}{2}}\sum_{p_1,p_2\in \mathbf{Z}} e(\tfrac{1}{2}\cdot(w_1(p_1)^2+w_2(p_2)^2)),$$

in which $(1-4w_1w_2)^{\frac{1}{2}}$ is positive when w_1, w_2 are pure imaginary. This is the formula we have liked to obtain.

Now, in the following, we agree to take (plus) one for $\det(c\tau+d)^{\frac{1}{2}}$ when the determinant is the constant one. Also, in determining the constant $\text{\ae}$, we shall use diagonal matrices for τ. We apply the transformation law of theta-functions to $M=A_{ij}$ for $i\neq j$. In the formula so obtained, we make all τ_{kk} except for $k=i, j$ tend to infinity along the imaginary axis. Then, if we replace τ_{ii}, τ_{jj} by w_1, w_2, we get the following formula

$$\sum_{p_1,p_2\in \mathbf{Z}} e(\tfrac{1}{2}\cdot((w_1+4w_2)(p_1)^2+w_2(p_2)^2+4w_2p_1p_2)) \\ =\text{\ae}(A_{ij})\sum_{p_1,p_2\in \mathbf{Z}} e(\tfrac{1}{2}\cdot(w_1(p_1)^2+w_2(p_2)^2)).$$

Since the binary quadratic form in the exponential on the left side is $w_1(p_1)^2+w_2(2p_1+p_2)^2$, we see that the two infinite series, which are nothing else than $\theta_0(w_1)\theta_0(w_2)$ with θ_0 for $g=1$, cancel each other. Therefore we have $\text{\ae}(A_{ij})=1$. In more or less the same way, we can determine $\text{\ae}(A_{ii})$ and $\text{\ae}(B_{ij})$. If we take $\det(c\tau+d)^{\frac{1}{2}}=e(\frac{1}{4})$ for $M=A_{ii}$, we have $\text{\ae}(A_{ii})=e(-\frac{1}{4})$ so that their product is one. Also we have $\text{\ae}(B_{ij})=1$ including the case $i=j$. For the calculation of $\text{\ae}(C_{ij})$, we made preparations already. In fact, if we follow the same method, we get

$$\sum_{p_1,p_2\in \mathbf{Z}} e(1/2(1-4w_1w_2)\cdot(w_1(p_1)^2+w_2(p_2)^2-4w_1w_2p_1p_2)) \\ =\text{\ae}(C_{ij})(1-4w_1w_2)^{\frac{1}{2}}\sum_{p_1,p_2\in \mathbf{Z}} e(\tfrac{1}{2}\cdot(w_1(p_1)^2+w_2(p_2)^2))$$

for $i<j$. Therefore, if $\det(c\tau+d)^{\frac{1}{2}}$ is positive when τ is a pure imaginary diagonal matrix, we have $\text{\ae}(C_{ij})=1$. Similarly, we have $\text{\ae}(C_{ii})=1$ if $\det(c\tau+d)^{\frac{1}{2}}$ is in the first quadrant when τ is a diagonal matrix. We shall summarize our results in the following way:

THEOREM 2. *With the choice of the square root* $\det(c\tau+d)^{\frac{1}{2}}$ *as explained, we have*

$$\text{\ae}(A_{ij})=1\ \ (i\neq j),\quad \text{\ae}(A_{ii})=e(-\tfrac{1}{4})$$
$$\text{\ae}(B_{ij})=\text{\ae}(C_{ij})=1\ \ (i\leq j).$$

This theorem allows us to obtain $\text{\ae}(M)$ for any given M in $\Gamma_g(2)$. We shall introduce a notation. If n is an integer, we shall denote by $(-4/n)$ the (value at n of the) Kronecker symbol for the prime discriminant -4. This means simply that $(-4/n)$ is 0 or ± 1 according as n is even or

$\pm 1 \bmod 4$. Now, if M is an element of $\Gamma_g(2)$ composed of a, b, c, d, we put

$$(-4/M) = \prod_{1 \leqq i, j \leqq g} (-4/d_{ij}).$$

We note that the product can be restricted to diagonal coefficients. Also we can as well take the product over the coefficients of a. The point is that $M \rightarrow (-4/M)$ is a character of $\Gamma_g(2)$ in the sense

$$(-4/MM') = (-4/M)(-4/M')$$

for M, M' in $\Gamma_g(2)$. Note also that we have $(-4/-M) = (-1)^g(-4/M)$. Now we shall consider products of theta-constants. Let $\mathfrak{m}$, $\mathfrak{n}$ be two even characteristics and put $\psi = \theta_{\mathfrak{m}}\theta_{\mathfrak{n}}$. We shall show that there exists a character χ of $\Gamma_g(2)$ satisfying

$$\psi(M \cdot \tau) = (-4/M)\chi(M)\det(c\tau + d)\psi(\tau)$$

for every M in $\Gamma_g(2)$. Since we have $M \cdot \mathfrak{m} \equiv \mathfrak{m} \bmod 2$ for every M in $\Gamma_g(2)$, the transformation law of theta-functions shows that we have such a relation with some constant $\chi(M)$, and clearly $\chi(-M) = \chi(M)$. On the other hand, if M and M' are arbitrary elements of $Sp(g, \boldsymbol{R})$ composed of a, b, c, d and a', b', c', d' respectively and if MM' is composed of a'', b'', c'', d'', we have $c''\tau + d'' = (c\tau' + d)(c'\tau + d')$ in which $\tau' = M' \cdot \tau$. The verification is straightforward. Therefore, by taking M and M' from $\Gamma_g(2)$, we get

$$\chi(MM') = \chi(M)\chi(M').$$

This shows that χ is a character of $\Gamma_g(2)/\pm 1_{2g}$. Now, using Theorem 2 we shall determine $\chi(M)$ for $M = A_{ij}, B_{ij}, C_{ij}$.

Since the computations are similar, we shall illustrate them in the case $M = B_{ii}$. We note first that the solution of $M \cdot \mathfrak{x} = \mathfrak{m}$ differs from $\mathfrak{m}$ only in its $(g + i)$-th coefficient, and this is $m_i'' + 2(m_i' - 1)$. Hence $\theta_{\mathfrak{x}}(\tau)$ and $\theta_{\mathfrak{m}}(\tau)$ differ by $(-1)^{m_i'(m_i'-1)} = 1$. Similarly we have $\theta_{\mathfrak{y}}(\tau) = \theta_{\mathfrak{n}}(\tau)$ for $M \cdot \mathfrak{y} = \mathfrak{n}$. Therefore, by the transformation law of theta-functions and by Theorem 2, we get

$$\begin{aligned}\psi(M \cdot \tau) &= \theta_{M \cdot \mathfrak{x}}(M \cdot \tau)\theta_{M \cdot \mathfrak{y}}(M \cdot \tau) \\ &= e(\phi_{\mathfrak{x}}(M) + \phi_{\mathfrak{y}}(M))\det(c\tau + d)\psi(\tau) \\ &= e(-\tfrac{1}{4} \cdot ((m_i')^2 + (n_i')^2 - 2m_i' - 2n_i')) \\ &\qquad \cdot \det(c\tau + d)\psi(\tau).\end{aligned}$$

Since we have $(-4/M) = 1$, this has to be $\chi(M)\det(c\tau + d)\psi(\tau)$. In this way, we get

$$\chi(B_{ii}) = (-1)^{m_i'+n_i'} e(-\tfrac{1}{4}\cdot((m_i')^2+(n_i')^2)).$$

The computations of other $\chi(M)$ are similar, and we have

$$\begin{aligned} \chi(A_{ij}) &= (-1)^{m_i'm_j''+n_i'n_j''} \\ \chi(B_{ij}) &= (-1)^{m_i'm_j'+n_i'n_j'} \qquad (i<j) \\ \chi(C_{ij}) &= (-1)^{m_i''m_j''+n_i''n_j''} \quad (i<j) \\ \chi(C_{ii}) &= e(-\tfrac{1}{4}\cdot((m_i'')^2+(n_i'')^2)). \end{aligned}$$

We shall formulate our results in the following way:

THEOREM 3. *Suppose that an element M of $\Gamma_g(2)$ is expressed as a product of powers of A_{ij}, B_{ij}, C_{ij} with exponents given by p, q, r and of an element of the commutator group of $\Gamma_g(2)$. Then $\psi=\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}$ satisfies the functional equation*

$$\psi(M\cdot\tau) = (-4/M)\chi(M)\det(c\tau+d)\psi(\tau)$$

in which

$$(-4/M) = (-1)^{\Sigma p_{ii}}, \qquad \chi(M) = (-1)^{A} e(-\tfrac{1}{4})^{B}$$

with

$$\begin{aligned} A = \sum_{1\leqq i,j\leqq g} p_{ij}(m_i'm_j''+n_i'n_j'') &+ \sum_{1\leqq i\leqq j\leqq g} q_{ij}(m_i'm_j'+n_i'n_j') \\ &+ \sum_{1\leqq i<j\leqq g} r_{ij}(m_i''m_j''+n_i''n_j'') \\ B = \sum_{1\leqq i\leqq g} q_{ii}((m_i')^2+(n_i')^2) &+ \sum_{1\leqq i\leqq g} r_{ii}((m_i'')^2+(n_i'')^2). \end{aligned}$$

In particular, if M is contained in the congruence group $\Gamma_g(4,8)$, we have $(-4/M)=1$ and $A\equiv 0 \bmod 2$, $B\equiv 0 \bmod 4$. Hence we have $\psi(M\cdot\tau) = \det(c\tau+d)\psi(\tau)$ for every M in $\Gamma_g(4,8)$. Also, in the case $g=1$, the necessary condition at infinity is satisfied for $\psi=\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}$. Therefore each $\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}$ is a modular form of weight 1 belonging to $\Gamma_g(4,8)$, i.e. it is an element of $A(\Gamma_g(4,8))_1$.

4. Vanishing of theta constants. We shall first recall some further properties of theta-functions (with brief indications of proofs) [cf. 2, 7, 8]. In general, suppose that G is a topological abelian group. Then a lattice in G is a discrete subgroup of G with compact factor group. Let Λ_1 and Λ_2 be two lattices in G which are commensurable (in the usual sense, i.e. their intersection has finite indices in both of them). Then $\Lambda=\Lambda_1+\Lambda_2$ is also a lattice in G. Suppose that f is a complex-valued function defined on G (or at least on Λ) for which $\sum f(n)$ is absolutely convergent when n runs over Λ. Then the summations of f over Λ_1 and Λ_2 are related as follows:

LEMMA 3. *We have*

$$\sum_{n_1} f(n_1) = 1/[\Lambda : \Lambda_1] \cdot \sum_{\chi, r} \Big(\sum_{n_2} \chi(n_2 + r) f(n_2 + r) \Big),$$

in which n_1, n_2 *run over* Λ_1, Λ_2 *while* χ *and* r *run over the dual of* Λ/Λ_1 *and over* Λ/Λ_2.

In fact, the lemma hardly needs any proof and it simply depends on the orthogonality of characters of finite abelian groups. However, it has a number of applications. Take, for instance, the group of four vectors in $\boldsymbol{R}^g$ as G and its subgroup of integer vectors as Λ_1. We consider elements of Λ_1 as $g \times 4$ matrices and multiply the following orthogonal symmetric matrix[2]

$$A = \tfrac{1}{2} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

to them from the right. Take the lattice of such elements as Λ_2, i. e. put $\Lambda_2 = \Lambda_1 A$. Then Λ_1 and Λ_2 permute by right multiplication of A and their index in Λ is 2^g. We shall use $(p \ 0 \ 0 \ 0)$ where the coefficients of p are 0, 1 as representatives of $\Lambda \bmod \Lambda_2$. As for the function f, let $\mathfrak{m}$, $\mathfrak{a}$, $\mathfrak{b}$ be three characteristics and put $\mathfrak{m}_1 = \mathfrak{m}$, $\mathfrak{m}_2 = \mathfrak{m} + \mathfrak{a}$, $\mathfrak{m}_3 = \mathfrak{m} + \mathfrak{b}$, $\mathfrak{m}_4 = \mathfrak{m} - \mathfrak{a} - \mathfrak{b}$. Further let τ be a point of $\mathfrak{S}_g$ and u_1, u_2, u_3, u_4 four points of $\boldsymbol{C}^g$. Put

$$\begin{aligned} f(x) = \prod_{1 \leqq \alpha \leqq 4} e\big(\tfrac{1}{2} \cdot {}^t(x_\alpha + m_\alpha'/2)\tau(x_\alpha + m_\alpha'/2) \\ + {}^t(x_\alpha + m_\alpha'/2)(u_\alpha + m_\alpha''/2)\big). \end{aligned}$$

We then apply Lemma 3 to these data and make a straightforward calculation. In this way we get

$$\begin{aligned} &\theta_{\mathfrak{m}}(\tau, u_1)\theta_{\mathfrak{m}+\mathfrak{a}}(\tau, u_2)\theta_{\mathfrak{m}+\mathfrak{b}}(\tau, u_3)\theta_{\mathfrak{m}-\mathfrak{a}-\mathfrak{b}}(\tau, u_4) \\ &\quad = 1/2^g \cdot \sum_{\mathfrak{n} \bmod 2} e(\mathfrak{m}, \mathfrak{n})\theta_{\mathfrak{n}}(\tau, v_1)\theta_{\mathfrak{n}+\mathfrak{a}}(\tau, v_2)\theta_{\mathfrak{n}+\mathfrak{b}}(\tau, v_3)\theta_{\mathfrak{n}-\mathfrak{a}-\mathfrak{b}}(\tau, v_4) \end{aligned}$$

in which $(u_1 \ u_2 \ u_3 \ u_4)$ and $(v_1 \ v_2 \ v_3 \ v_4)$ are related by right multiplications of A. We observe that A is nothing else than the case $g = 1$ of the $2^{2g} \times 2^{2g}$ coefficient matrix $(1/2^g \cdot e(\mathfrak{m}, \mathfrak{n}))$, which is orthogonal symmetric (and proper for $g \geqq 2$). At any rate, the above formula is known as "Riemann's theta-formula" while its full meaning was known to Jacobi in the case $g = 1$.

---

[2] We note that any improper orthogonal symmetric matrix A^* (of degree four) such that $2A^*$ but not A^* itself has integer coefficients can be transformed into $\pm A$ by an integer orthogonal matrix.

We make another application of Lemma 3 which we shall use in this section. We take the group of two vectors in $\boldsymbol{R}^g$ as G and its subgroup of integer vectors as Λ_1. This time we multiply the following matrix

$$A = \tfrac{1}{2}\cdot\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

to Λ_1 from the right to define Λ_2. Then we have $\Lambda = \Lambda_2$ and the index of Λ_1 in Λ is 2^g. As for the function f, let $\mathfrak{m}_1$, $\mathfrak{m}_2$ be two characteristics with $m_1'' = m_2'' = m''$ and put

$$f(x) = \prod_{\alpha=1,2} e\bigl({}^t(x_\alpha + m_\alpha'/2)\tau(x_\alpha + m_\alpha'/2) \\ + {}^t(x_\alpha + m_\alpha'/2)(2u_\alpha + m''/2)\bigr).$$

We then apply Lemma 3 to these data and immediately get

$$\theta\left(\begin{smallmatrix} m_1' \\ m'' \end{smallmatrix}\right)(2\tau, 2u_1)\,\theta\left(\begin{smallmatrix} m_2' \\ m'' \end{smallmatrix}\right)(2\tau, 2u_2) = \\ 1/2^g\cdot\sum_{n''\,\mathrm{mod}\,2}(-1)^{{}^t m_1' n''}\theta\begin{smallmatrix} m_1'+m_2' \\ m''+n'' \end{smallmatrix}\Big)(\tau, u_1+u_2)\,\theta\left(\begin{smallmatrix} m_1'+m_2' \\ n'' \end{smallmatrix}\right)(\tau, u_1-u_2).$$

We multiply $(-1)^{{}^t m_1' r''}$ to both sides and take a summation over $m_1'\,\mathrm{mod}\,2$. In this way we can solve the system of equations and we get

$$\theta\left(\begin{smallmatrix} m' \\ m''+r'' \end{smallmatrix}\right)(\tau, u_1+u_2)\,\theta\left(\begin{smallmatrix} m' \\ r'' \end{smallmatrix}\right)(\tau, u_1-u_2) \\ = \sum_{m_1'\,\mathrm{mod}\,2}(-1)^{{}^t m_1' r''}\theta\left(\begin{smallmatrix} m_1' \\ m'' \end{smallmatrix}\right)(2\tau, 2u_1)\,\theta\left(\begin{smallmatrix} m'+m_1' \\ m'' \end{smallmatrix}\right)(2\tau, 2u_2).$$

If we put $m_1' = m_2'$, $u_2 = 0$ in the first formula and $m'' = 0$, $u_1 = u_2$ in the second formula, changing the notation slightly, we have

$$\theta_{\mathfrak{m}}(\tau)\theta_{\mathfrak{m}}(\tau, 2z) = 1/2^g\cdot\sum_{n''\,\mathrm{mod}\,2}(-1)^{{}^t m' n''} \\ \theta\left(\begin{smallmatrix} 0 \\ m''+n'' \end{smallmatrix}\right)(\tau/2, z)\,\theta\left(\begin{smallmatrix} 0 \\ n'' \end{smallmatrix}\right)(\tau/2, z)$$

$$\theta_{\mathfrak{m}}(\tau)\theta_{\mathfrak{m}}(\tau, z) = \sum_{n'\,\mathrm{mod}\,2}(-1)^{{}^t m'' n'}\theta\left(\begin{smallmatrix} m'+n' \\ 0 \end{smallmatrix}\right)(2\tau, z)\,\theta\left(\begin{smallmatrix} n' \\ 0 \end{smallmatrix}\right)(2\tau, z).$$

We call them the *first* and the *second* (principal) *transformations of degree two.*

Now we shall start using these formulas. We shall prove the following lemma:

LEMMA 4. *Suppose that τ is any given point of $\mathfrak{S}_g$. Then at least one theta-constant does not vanish at τ.*

Proof. If the statement is not true, we have $\theta_{\mathfrak{m}}(\tau) = 0$ for all $\mathfrak{m}$ at some point τ of $\mathfrak{S}_g$. We can then derive a contradiction using either one of the transformations of degree two. For instance, the second transformation implies

$$\sum_{n' \bmod 2} (-1)^{{}^t m'' n'} \theta\binom{n'}{0}(2\tau, z)^2 = 0$$

for all m''. By solving this system of 2^g equations, we get

$$\theta\binom{n'}{0}(2\tau, z) = 0.$$

But no theta-function vanishes identically in z (no matter what τ or in this case 2τ is). This proves the lemma.

We shall see later that Lemma 4 has an important consequence. It is rather surprising that its clear statement does not exist in any well-known literature we came across. We shall add further comments to Lemma 4.

In general, we consider a subgroup $\mathfrak{g}$ of the additive group of characteristics mod 2 such that $e(\mathfrak{m}, \mathfrak{n}) = 1$ for all $\mathfrak{m}$, $\mathfrak{n}$ in $\mathfrak{g}$. Any maximal subgroup with this property is called a Göpel group. Clearly, the group of characteristics mod 2 satisfying $m' \equiv 0 \bmod 2$ or $m'' \equiv 0 \bmod 2$ is a Göpel group and Göpel groups are of order 2^g. Also, if $\mathfrak{g}$ is a Göpel group, it has 2^g cosets and they are called Göpel systems mod $\mathfrak{g}$. Now the finite group $Sp(g, \mathbf{Z}/2\mathbf{Z}) = \Gamma_g(1)/\Gamma_g(2)$ is precisely the orthogonal group in the vector space of characteristics mod 2 with respect to the quadratic form $e(\mathfrak{m}, \mathfrak{n})$ (when the multiplication of elements of $Sp(g, \mathbf{Z}/2\mathbf{Z})$ and vectors is an ordinary matrix multiplication). Therefore, if $Sp(g, \mathbf{Z}/2\mathbf{Z})$ operates on the vector space as $\mathfrak{m} \to M \cdot \mathfrak{m} \bmod 2$, Göpel systems are mapped to Göpel systems. Moreover the stabilizer of the Göpel group defined by $m'' \equiv 0 \bmod 2$ is the subgroup of $Sp(g, \mathbf{Z}/2\mathbf{Z})$ defined by $b \equiv 0 \bmod 2$. Since the order of this subgroup is equal to the number of symmetric matrices mod 2 of degree g multiplied by the order of $SL(g, \mathbf{Z}/2\mathbf{Z})$, its index in $Sp(g, \mathbf{Z}/2\mathbf{Z})$ is $(2^g + 1)(2^{g-1} + 1) \cdots (2 + 1)$. Since this is clearly the total number of Göpel groups, there exist as many Göpel systems consisting of even characteristics. On the other hand, if a Göpel group $\mathfrak{g}$ is given, there exists one and only one Göpel system mod $\mathfrak{g}$ which consists of same kind of characteristics. In fact, because of

$$e(\mathfrak{m} + \mathfrak{n}) = e(\mathfrak{m}) e(\mathfrak{n}) e(\mathfrak{m}, \mathfrak{n}),$$

the function $\mathfrak{m} \to e(\mathfrak{m})$ is a character of $\mathfrak{g}$. However any character of $\mathfrak{g}$ can be written in the form $e(\mathfrak{m}, \mathfrak{a})$ with some characteristic $\mathfrak{a}$ uniquely determined mod $\mathfrak{g}$. Then the Göpel system $\mathfrak{g} + \mathfrak{a}$ consists of characteristics of the same kind while other Göpel systems mod $\mathfrak{g}$ are mixed. Therefore, we see that the Göpel system $\mathfrak{g} + \mathfrak{a}$ consists of even characteristics and the total number of even characteristics is $2^{g-1}(2^g+1)$ while the total number of odd characteristics is $2^{g-1}(2^g-1)$. Furthermore $Sp(g,\boldsymbol{Z}/2\boldsymbol{Z})$ operates transitively on the set of Göpel systems consisting of even characteristics.

Now, suppose that $\mathfrak{g}$ is an arbitrary Göpel system. Then, for any given point τ of $\mathfrak{S}_g$, we shall denote by $n_{\mathfrak{g}}(\tau)$ the number of linearly independent functions (in z of course) among $\theta_{\mathfrak{m}}(\tau,z)\theta_{\mathfrak{n}}(\tau,z)$ over $\boldsymbol{C}$ with $\mathfrak{m}$, $\mathfrak{n}$ in $\mathfrak{g}$. Then, of we increase z by $\frac{1}{2}\cdot(\tau 1_g)\mathfrak{a}$, we see that $n_{\mathfrak{g}+\mathfrak{a}}(\tau)=n_{\mathfrak{g}}(\tau)$ for every $\mathfrak{a}$. Also the transformation law of theta-functions implies

$$n_{M\cdot\mathfrak{g}}(M\cdot\tau)=n_{\mathfrak{g}}(\tau)$$

for every M in $\Gamma_g(1)$. Therefore we will know $n_{\mathfrak{g}}(\tau)$ in general if we know it for any one of the Göpel systems. We take the Göpel group defined by $m'\equiv 0 \bmod 2$ as $\mathfrak{g}$. Then we have the following remarkable relation

$$n_{\mathfrak{g}}(\tau)=\textit{number of non-vanishing } \theta_{\mathfrak{a}}(2\tau) \textit{ for all } \mathfrak{a}.$$

We first observe that the 2^{2g} theta-functions $\theta_{\mathfrak{a}}(\tau,z)$ are linearly independent over $\boldsymbol{C}$ for any given τ. Otherwise, suppose that there exists a linear relation between $\theta_{\mathfrak{a}}(\tau,z)$ with complex coefficients $c_{\mathfrak{a}}$ say. Increase z by $(\tau 1_g)\mathfrak{n}$ and use the functional equation. Then we get $c_{\mathfrak{a}}\theta_{\mathfrak{a}}(\tau,z)=0$,hence $c_{\mathfrak{a}}=0$ for all $\mathfrak{a}$. Now the first transformation of degree two shows that the (complex) vector space of functions generated (as linear combinations) by $\theta_{\mathfrak{a}}(2\tau)\theta_{\mathfrak{a}}(2\tau,2z)$ coincides with the vector space of functions generated by $\theta_{\mathfrak{m}}(\tau,z)\theta_{\mathfrak{n}}(\tau,z)$ with $m'\equiv n'\equiv 0 \bmod 2$. Accordingly, we can compute the dimension of the vector space in two ways and we get the said relation. We note that this was at the basis of the proof of Lemma 4. As an extremely special case, we see that, in the case $g=1$, none of the three theta-constants $\theta_{00}(w)$, $\theta_{01}(w)$, $\theta_{10}(w)$ vanishes at any point w in the upper-half plane. This fact will be used in the next section.

5. Field of modular functions. In general, if Γ is a discrete subgroup of $Sp(g,\boldsymbol{R})$ commensurable with $\Gamma_g(1)$, we shall denote by $F(\Gamma)$ the field of homogeneous elements of degree zero in the field of fractions of $A(\Gamma)$ and call it the *field of modular functions* belonging to Γ. This is an algebraic function field of dimension $\frac{1}{2}\cdot g(g+1)$ over $\boldsymbol{C}$. The purpose of this section is to prove the following theorem:

THEOREM 4. *The field generated over* $\boldsymbol{C}$ *by quotients of theta-constants contains* $F(\Gamma_g(1))$ *as its subfield.*

The proof of this important theorem consists of several steps most of which are due to Wirtinger [12]. However, since certain arguments (see

e. g., p. 29) are not too convincing, we shall make some modifications. Suppose that τ is a point of $\mathfrak{S}_g$, z a point of $\boldsymbol{C}^g$ and k a non-negative integer. Then an analytic function $\Theta(z)$ of z satisfying the functional equation

$$\Theta(z+(\tau 1_g)\mathfrak{n})=e(\mathfrak{m},\mathfrak{n})e(-k({}^tn'z+\tfrac{1}{2}\cdot{}^tn'\tau n'))\Theta(z)$$

for every $\mathfrak{n}$ in $\boldsymbol{Z}^{2g}$ is called a theta-function of order k and of characteristic $\mathfrak{m}$. The set of such functions forms a complex vector space and, by using Fourier expansions of $e(-\frac{1}{2}\cdot{}^tm'z)\Theta(z)$ as in the case $k=1$, we can show that its dimension is equal to k^g. We shall denote by S_k the vector space of theta-functions of order $4k$ and of characteristic zero, and consider the graded ring

$$S=\bigoplus_{0\leqq k<\infty} S_k$$

over $S_0=\boldsymbol{C}$. Clearly $\theta_{\mathfrak{m}}(\tau,2z)$ are elements of S_1 (as functions of z) and they are linearly independent. Therefore the 2^{2g} theta-functions $\theta_{\mathfrak{m}}(\tau,2z)$ form a base of S_1. We shall show that they generate S at least when τ is contained in some "admissible set" by which we understand the complement of a countable union of closed analytic subsets of $\mathfrak{S}_g$ having empty intersections with the set of diagonal matrices.

We have only to show that among all k-fold products of $\theta_{\mathfrak{m}}(\tau,2z)$ there exist $(4k)^g$ linearly independent ones. For this purpose, we can replace $2z$ by z. Since the linear independence can be determined by the rank of the (infinite) coefficient matrix of their Fourier expansions, we have only to construct $(4k)^g$ products of $\theta_{\mathfrak{m}}(\tau,z)$ which are linearly independent when τ is a diagonal matrix. We say that a k-fold product of $\theta_{\mathfrak{m}}(\tau,z)$ for the characteristics $\mathfrak{m}_1,\cdots,\mathfrak{m}_k$ is canonical if, in the $2g\times k$ matrix $(\mathfrak{m}_1\cdots\mathfrak{m}_k)$, the $2\times k$ submatrix composed of i-th and $(g+i)$-th lines is one of the following four types

$$\begin{pmatrix}0\cdots0 & 0\cdots0\\ 0\cdots0 & 1\cdots1\end{pmatrix}\qquad\begin{pmatrix}0\cdots0 & 0\cdots0 & 1\\ 0\cdots0 & 1\cdots1 & 0\end{pmatrix}$$

$$\begin{pmatrix}0\cdots0 & 0\cdots0 & 1\\ 0\cdots0 & 1\cdots1 & 1\end{pmatrix}\qquad\begin{pmatrix}0\cdots0 & 0\cdots0 & 1 & 1\\ 0\cdots0 & 1\cdots1 & 0 & 1\end{pmatrix}$$

for $i=1,\cdots,g$. The number of canonical products is clearly $(4k)^g$. On the other hand, the following $4k$ products of elliptic theta-functions

$$\begin{array}{ll}\theta_{00}(w,z)^p\theta_{01}(w,z)^q & p+q=k\\ \theta_{00}(w,z)^p\theta_{01}(w,z)^q\theta_{10}(w,z) & p+q=k-1\\ \theta_{00}(w,z)^p\theta_{01}(w,z)^q\theta_{11}(w,z) & p+q=k-1\\ \theta_{00}(w,z)^p\theta_{01}(w,z)^q\theta_{10}(w,z)\theta_{11}(w,z) & p+q=k-2\end{array}$$

are linearly independent for every w in the upper-half plane. In fact $\theta_{00}(w, z)$, $\theta_{01}(w, z)$, $\theta_{10}(w, z)$ are even while $\theta_{11}(w, z)$ is odd. Moreover $\theta_{00}(w, z)$ and $\theta_{01}(w, z)$ are invariant while $\theta_{10}(w, z)$ changes its sign by $z \to z + 1$. Therefore we have only to show that the first $k + 1$ products are linearly independent. But this is a consequence of the fact that $\theta_{00}(w, z)$ and $\theta_{01}(w, z)$ are linearly independent. Therefore the $(4k)^g$ canonical products are linearly independent when τ is a diagonal matrix. This proves the assertion.

As a consequence, if τ is in the admissible set, we have

$$\theta_{\mathfrak{n}_1}(\tau, z) \cdots \theta_{\mathfrak{n}_k}(\tau, z) = \sum_{\text{canonical}} A_{\mathfrak{m}_1 \cdots \mathfrak{m}_k} \theta_{\mathfrak{m}_1}(\tau, z) \cdots \theta_{\mathfrak{m}_k}(\tau, z)$$

with uniquely determined coefficients $A_{\mathfrak{m}_1 \cdots \mathfrak{m}_k}$ for every $\mathfrak{n}_1, \cdots, \mathfrak{n}_k$. The summation is extended over the $(4k)^g$ canonical products. We shall investigate the coefficients in the case $k = 2$. Let $\theta_{\mathfrak{n}_1}(\tau, z)\theta_{\mathfrak{n}_2}(\tau, z)$ be an arbitrary product. Then there exist 2^g canonical products $\theta_{\mathfrak{m}_1}(\tau, z)\theta_{\mathfrak{m}_2}(\tau, z)$ satisfying $\mathfrak{m}_1 + \mathfrak{m}_2 \equiv \mathfrak{n}_1 + \mathfrak{n}_2 \bmod 2$ and $\theta_{\mathfrak{n}_1}(\tau, z)\theta_{\mathfrak{n}_2}(\tau, z)$ is a linear combination of them. If we replace z in this relation by $\frac{1}{2} \cdot (\tau 1_g)\mathfrak{n}$, we get

$$\theta_{\mathfrak{n}_1+\mathfrak{n}}(\tau)\ \theta_{\mathfrak{n}_2+\mathfrak{n}}(\tau) = \sum A_{\mathfrak{m}_1\mathfrak{m}_2} e\left(\tfrac{1}{4} \cdot {}^t n'(n_1'' + n_2'' - m_1'' - m_2'')\right) \cdot \theta_{\mathfrak{m}_1+\mathfrak{n}}(\tau)\theta_{\mathfrak{m}_2+\mathfrak{n}}(\tau).$$

We shall show that the $2^g \times 2^{2g}$ matrix $(\pm \theta_{\mathfrak{m}_1+\mathfrak{n}}(\tau)\theta_{\mathfrak{m}_2+\mathfrak{n}}(\tau))$ has rank 2^g when τ is a diagonal matrix, hence when τ is contained in some admissible set. If τ is a diagonal matrix, the problem is reduced immediately to the case $g = 1$. In this case the matrix in question has rank two because none of the three even theta-constants vanishes at any point of the upper-half plane. Because of this, the $A_{\mathfrak{m}_1\mathfrak{m}_2}$ can be determined by solving the above system of linear equations. Hence they are contained in the field generated over $\boldsymbol{Q}$ by $\theta_{\mathfrak{m}_1}(\tau)\theta_{\mathfrak{n}_1}(\tau)/\theta_{\mathfrak{m}_2}(\tau)\theta_{\mathfrak{n}_2}(\tau)$ where the four characteristics are even and satisfy $\mathfrak{m}_1 + \mathfrak{n}_1 \equiv \mathfrak{m}_2 + \mathfrak{n}_2 \bmod 2$. It is assumed, thereby, that τ is in some admissible set.

Now pick a point τ_0 of $\mathfrak{S}_g$ which is a diagonal matrix with diagonal coefficients $w_1, \cdots, w_g$. Also we shall denote by $\tau = (\tau_{ij})$ a variable matrix. Then the set of convergent power-series in τ_{ij} for $i \neq j$ and $\tau_{ii} - w_i$ forms a regular local ring R. The previous argument shows that the $A_{\mathfrak{m}_1 \cdots \mathfrak{m}_k}$ are contained in R. We associate a letter $X_{\mathfrak{m}}$ to each characteristic $\mathfrak{m}$ with coefficients 0, 1 and consider the ring of polynomials $R[X]$ in the set X of 2^{2g} letters $X_{\mathfrak{m}}$ with coefficients in R. Also certain monomials in $X_{\mathfrak{m}}$ are called canonical in an obvious way. We shall denote the field of fractions of R by F. If $f = f(\tau)$ is an element of R, the correspondence $f \to f(\tau_0)$ gives the canonical homomorphism of R to its residue field $\boldsymbol{C}$. Moreover, if we replace $X_{\mathfrak{m}}$ by $\theta_{\mathfrak{m}}(\tau, z)$, we get a homomorphism from $R[X]$ to the ring of polynomials in $\theta_{\mathfrak{m}}(\tau, z)$ with coefficients in R. We shall denote its kernel by I. Then I is a homogeneous ideal of $R[X]$ and the homogeneous ideal J of $R[X]$ generated by

$$X_{\mathfrak{n}_1}X_{\mathfrak{n}_2} - \sum_{\text{canonical}} A_{\mathfrak{m}_1\mathfrak{m}_2} X_{\mathfrak{m}_1}X_{\mathfrak{m}_2}$$

is contained in I. Our purpose is to prove $F[X]J = F[X]I$. We shall first show that the image of J under the canonical homomorphism $R[X] \rightarrow \boldsymbol{C}[X]$, which we shall denote by $\boldsymbol{C}[X]J$, coincides with the kernel of the homomorphism from $\boldsymbol{C}[X]$ to the ring of polynomials in $\theta_{\mathfrak{m}}(\tau_0, z)$ with coefficients in $\boldsymbol{C}$ (defined by the substitution of $X_{\mathfrak{m}}$ by $\theta_{\mathfrak{m}}(\tau_0, z)$). We observe that in the elliptic case there exist two non-canonical products and they are $\theta_{10}(w, z)^2$ and $\theta_{11}(w, z)^2$. Furthermore they are linear combinations of $\theta_{00}(w, z)^2$ and $\theta_{01}(w, z)^2$. Suppose that $P(w, z) = Q(w, z)$ is either one of these two relations and replace (w, z) by (w_i, z_i) for any i satisfying $1 \leqq i \leqq g$. Choose $g-1$ canonical products out of eight and replace (w, z) by $(w_1, z_1), \cdots, (w_{i-1}, z_{i-1}), (w_{i+1}, z_{i+1}), \cdots, (w_g, z_g)$, and multiply to $P(w_i, z_i) = Q(w_i, z_i)$. In this way we get a relation of the form

$$\theta_{\mathfrak{n}_1}(\tau_0, z)\theta_{\mathfrak{n}_2}(\tau_0, z) = \sum_{\text{canonical}} A_{\mathfrak{m}_1\mathfrak{m}_2}'\theta_{\mathfrak{m}_1}(\tau_0, z)\theta_{\mathfrak{m}_2}(\tau_0, z)$$

with $A_{\mathfrak{m}_1\mathfrak{m}_2}'$ in $\boldsymbol{C}$. Because of the linear independence of the canonical products, if $A_{\mathfrak{m}_1\mathfrak{m}_2} = A_{\mathfrak{m}_1\mathfrak{m}_2}(\tau)$ correspond to $\theta_{\mathfrak{n}_1}(\tau, z)\theta_{\mathfrak{n}_2}(\tau, z)$, we have $A_{\mathfrak{m}_1\mathfrak{m}_2}' = A_{\mathfrak{m}_1\mathfrak{m}_2}(\tau_0)$. Hence

$$X_{\mathfrak{n}_1}X_{\mathfrak{n}_2} - \sum_{\text{canonical}} A_{\mathfrak{m}_1\mathfrak{m}_2}'X_{\mathfrak{m}_1}X_{\mathfrak{m}_2}$$

is contained in $\boldsymbol{C}[X]J$. We observe that $A_{\mathfrak{m}_1\mathfrak{m}_2}'$ do not change even if we permute j-th and $(g+j)$-th coefficients of $\mathfrak{n}_1$ with those of $\mathfrak{n}_2$ for $j = 1, 2, \cdots, g$. Therefore, if $\mathfrak{n}_1$, $\mathfrak{n}_2$ become $\mathfrak{n}_1'$, $\mathfrak{n}_2'$ by this permutation, then

$$X_{\mathfrak{n}_1}X_{\mathfrak{n}_2} - X_{\mathfrak{n}_1'}X_{\mathfrak{n}_2'}$$

is also contained in $\boldsymbol{C}[X]J$. It is then clear that any homogeneous polynomial in $\boldsymbol{C}[X]$ is congruent to a linear combination of the canonical products of the same degree mod $\boldsymbol{C}[X]J$. Therefore, if the polynomial vanishes by the substitution $X_{\mathfrak{m}} \rightarrow \theta_{\mathfrak{m}}(\tau_0, z)$, it is contained in $\boldsymbol{C}[X]J$, and this is what we wanted to prove. We shall now recall the following simple lemma:

LEMMA 5. *Let R be an integral domain, F its field of fractions and C the field of fractions of a residue ring of R modulo one of its prime ideals.*

Let X be a finite set of letters and J a homogeneous ideal of $R[X]$. Then we have

$$\dim_{C}(C[X]/C[X]J)_k \geqq \dim_F(F[X]/F[X]J)_k$$

in which the subscript k indicates the homogeneous parts of degree k.

Proof. We have only to prove

$$\dim_C(C[X]J)_k \leqq \dim_F(F[X]J)_k.$$

Suppose that $p_1(X),\cdots,p_n(X)$ form a base of $(C[X]J)_k$ over C. We can assume that they are images of $P_1(X),\cdots,P_n(X)$ in J_k under the canonical homomorphism $R[X]\to C[X]$. We arrange monomials in X of degree k in a sequence as $Y_1,\cdots,Y_N$ and write

$$P_i(X)=\sum_{1\leqq j\leqq N} c_{ij}Y_j \qquad\qquad 1\leqq i\leqq n$$

with c_{ij} in R. Let γ_{ij} be the images of c_{ij} under the canonical homomorphism $R\to C$. Then we have $\operatorname{rank}(\gamma_{ij})=n$, hence $\operatorname{rank}(c_{ij})=n$. Consequently $P_1(X),\cdots,P_n(X)$ are linearly independent over F. This proves the lemma.

If we apply this lemma to the present case (with $C=\boldsymbol{C}$), observing

$$\dim_{\boldsymbol{C}}(\boldsymbol{C}[X]/\boldsymbol{C}[X]J)_k=\dim_F(F[X]/F[X]I)_k=(4k)^g$$

and the fact that J is contained in I, we get $(F[X]I)_k=(F[X]J)_k$ for $k=0,1,\cdots$. Since $F[X]I$ and $F[X]J$ are homogeneous ideals, this implies $F[X]I=F[X]J$.

Going back to the notation where τ means a point of $\mathfrak{S}_g$, we see that the mapping $z\to(\theta_{\mathfrak{m}}(\tau,2z))$ defines an analytic isomorphism of the complex torus $\boldsymbol{T}=\boldsymbol{C}^g/(\tau 1_g)\boldsymbol{Z}^{2g}$ to an abelian variety in a complex projective space of dimension $2^{2g}-1$, which is nothing else than the projective variety associated with the graded ring S. Furthermore, the defining ideal of this projective variety has a basis consisting of homogeneous polynomials of degree two with coefficients in $\boldsymbol{Q}(\theta_{\mathfrak{m}_1}(\tau)\theta_{\mathfrak{n}_1}(\tau)/\theta_{\mathfrak{m}_2}(\tau)\theta_{\mathfrak{n}_2}(\tau))$ where the characteristics are even and satisfy $\mathfrak{m}_1+\mathfrak{n}_1\equiv\mathfrak{m}_2+\mathfrak{n}_2 \bmod 2$. It is assumed that τ is in some admissible set. Therefore $\boldsymbol{Q}(\theta_{\mathfrak{m}_1}\theta_{\mathfrak{n}_1}/\theta_{\mathfrak{m}_2}\theta_{\mathfrak{n}_2})$ is a field of definition in the sense of Weil [11] of a generic abelian variety of the family and it contains the "field of moduli." It is known [cf. 1-20] that, if we extend the constant field from $\boldsymbol{Q}$ to $\boldsymbol{C}$, the field of moduli becomes $F(\Gamma_g(1))$. Therefore $\boldsymbol{C}(\theta_{\mathfrak{m}_1}\theta_{\mathfrak{n}_1}/\theta_{\mathfrak{m}_2}\theta_{\mathfrak{n}_2})$ contains $F(\Gamma_g(1))$. This completes the proof of Theorem 4.

6. Fundamental theorem. We shall combine the preceding results to obtain the "fundamental lemma" stated in the introduction. We first make some simple observations. Suppose that U is a complex manifold on which a group G operates as a group of analytic automorphisms. A non-vanishing analytic function R defined on U and depending on M in G is called a factor of automorphy (for G) if it satisfies $R(M_1M_2,\tau)=R(M_1,M_2\cdot\tau)R(M_2,\tau)$ for all M_1, M_2 in G. If M is an arbitrary element of G and f an analytic function in U, clearly

$$(M\cdot f)(\tau)=R(M^{-1},\tau)^{-1}f(M^{-1}\cdot\tau)$$

is an analytic function in U, and we have

$$M_1\cdot(M_2\cdot f)=(M_1M_2)\cdot f,\quad \text{identity}\cdot f=f.$$

Therefore G operates on the ring of analytic functions in U. Now suppose that Γ is a (discrete) subgroup of G. Then an analytic function f in U is called Γ-automorphic (for R) if it satisfies $M\cdot f=f$ for all M in Γ. We see immediately that, if f is Γ-automorphic, then $M\cdot f$ is $M\Gamma M^{-1}$-automorphic. There are two special cases which are of some importance. Suppose that Γ is contained in another subgroup Γ^* of G as a normal subgroup. Then Γ^* operates on the complex vector space of Γ-automorphic forms as above (and invariant elements are precisely Γ^*-automorphic forms). Another case is the following. Suppose that U is the Siegel upper-half plane $\mathfrak{S}_g$ and T an element of $Sp(g,\mathbf{R})$ such that $m^{\frac{1}{2}}T$ has integer coefficients for some positive integer m. Then $T\Gamma_g(n)T^{-1}$ contains $\Gamma_g(mn)$. Hence, if f is $\Gamma_g(n)$-automorphic, then $T\cdot f$ is $\Gamma_g(mn)$-automorphic. Also, in the case $m=2$ and n even, we see that $T\Gamma_g(n,2n)T^{-1}$ contains $\Gamma_g(2n,4n)$. Hence, if f is $\Gamma_g(n,2n)$-automorphic, then $T\cdot f$ is $\Gamma_g(2n,4n)$-automorphic.

Now, if M is an element of $Sp(g,\mathbf{R})$ composed of a, b, c, d, then

$$R(M,\tau)=\det(c\tau+d)^k$$

defines a factor of automorphy for any integer k. Moreover, if Γ is a subgroup of $Sp(g,\mathbf{R})$ commensurable with $\Gamma_g(1)$, modular forms of weight k belonging to Γ are (except for a modification in the case $g=1$) precisely Γ-automorphic forms for this R. Therefore the general consideration can be applied to our case. In this way, for instance, if Γ is a normal subgroup of $\Gamma_g(1)$, the finite group $\Gamma_g(1)/\Gamma$ operates on the graded ring $A(\Gamma)$ and we can try to decompose each homogeneous component of $A(\Gamma)$ into irreducible stable subspaces or, at least, determine the character as Hecke (and his student) did in the case $g=1$ for $\Gamma=\Gamma_1(q)$, q a prime number. However, in order to be able to do this, we have to have sufficient informations about the ring $A(\Gamma)$ (and about the group $\Gamma_g(1)/\Gamma$). For the congruence group

$\Gamma_g(4,8)$ (defined by $M \equiv 1_{2g} \bmod 4$, $(b)_0 \equiv (c)_0 \equiv 0 \bmod 8$), the following theorem is fundamental:

THEOREM 5. *The ring* $A(\Gamma_g(4,8))$ *of modular forms belonging to* $\Gamma_g(4,8)$ *is the normalization of* $\boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}]$ *in its field of fractions.*

We need the following lemma:

LEMMA 6. *Let* $\psi_1, \cdots, \psi_n$ *be modular forms belonging to* Γ *with no common zero in the quotient variety* $\Gamma\backslash\mathfrak{S}_g$ *and its boundary. Then the ring* $A(\Gamma)$ *is integral over* $\boldsymbol{C}[\psi_1, \cdots, \psi_n]$.

In fact, the theorem of compactifications reduces the problem to a purely algebraic statement. This algebraic statement is a consequence of Hilbert's "Nullstellensatz." Since Hilbert's original treatment is simple and elegant, we refer the proof to his paper [3].

We shall apply the lemma in the following way. Let $\mathfrak{m}'$ be an even characteristic for genus $g-1$ and add two more coefficients $m_g' = 0$, m_g'' = arbitrary to get an even characteristic $\mathfrak{m}$ (for genus g). Then, for every point τ' of $\mathfrak{S}_{g-1}$, we have

$$(\Phi \cdot \theta_{\mathfrak{m}})(\tau') = \lim_{\eta \to \infty} \theta_{\mathfrak{m}} \begin{pmatrix} \tau' & 0 \\ 0 & i\eta \end{pmatrix} = \theta_{\mathfrak{m}'}(\tau').$$

Consequently, the "Siegel operator" Φ applied to the ring of theta-constants of genus g gives a *surjection* to the ring of theta-constants of genus $g-1$. Since $\Gamma_g(1)$ operates on the ring $\boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}]$ of theta-constants, this combined with Lemma 4 shows that the modular forms $\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}$ belonging to $\Gamma_g(4,8)$ have no common zero in the quotient variety $\Gamma_g(4,8)\backslash\mathfrak{S}_g$ and its boundary. Therefore we can apply Lemma 6 and conclude that $A(\Gamma_g(4,8))$ is integral over $\boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}]$. On the other hand, we know by Theorem 4 that the subfield $\boldsymbol{C}(\theta_{\mathfrak{m}}/\theta_{\mathfrak{n}})$ of $F(\Gamma_g(4,8))$ contains $F(\Gamma_g(1))$. Since $F(\Gamma_g(4,8))$ is a finite Galois extension of $F(\Gamma_g(1))$, we can apply the theory of Galois to determine $\boldsymbol{C}(\theta_{\mathfrak{m}}/\theta_{\mathfrak{n}})$. We recall that, if M is an element of $\Gamma_g(1)$, it gives rise to an automorphism of the graded ring $A(\Gamma_g(4,8))$, hence of the field $F(\Gamma_g(4,8))$ as we explained. In this way, we get an epimorphism of $\Gamma_g(1)$ to the Galois group of $F(\Gamma_g(4,8))$ over $F(\Gamma_g(1))$ with $\Gamma_g(4,8)(\pm 1_{2g})$ as its kernel. We shall examine the condition for M to keep $\boldsymbol{C}(\theta_{\mathfrak{m}}/\theta_{\mathfrak{n}})$ elementwise invariant. The condition is $\theta_{\mathfrak{m}}(M^{-1}\cdot\tau)/\theta_{\mathfrak{n}}(M^{-1}\cdot\tau) = \theta_{\mathfrak{m}}(\tau)/\theta_{\mathfrak{n}}(\tau)$ for all even $\mathfrak{m}$, $\mathfrak{n}$ and, replacing τ by $M\cdot\tau$, we have the same condition with M^{-1} replaced by M. If we apply the transformation law of theta-functions to this condition, we see that $\theta_{\mathfrak{m}}(\tau)\theta_{M\cdot\mathfrak{n}}(\tau)$ and $\theta_{\mathfrak{n}}(\tau)\theta_{M\cdot\mathfrak{m}}(\tau)$ differ by a constant factor. We apply B_{ii} and C_{ii} to τ in $(\theta_{\mathfrak{m}}(\tau)\theta_{M\cdot\mathfrak{n}}(\tau))^2 = \text{const.}\ (\theta_{\mathfrak{n}}(\tau)\theta_{M\cdot\mathfrak{m}}(\tau))^2$ and we get $\mathfrak{m}+M\cdot\mathfrak{n}\equiv\mathfrak{n}+M\cdot\mathfrak{m} \bmod 2$, hence ${}^tM^{-1}(\mathfrak{m}-\mathfrak{n})\equiv\mathfrak{m}-\mathfrak{n} \bmod 2$ for all even $\mathfrak{m}$, $\mathfrak{n}$. This implies $M\equiv 1_{2g} \bmod 2$. Therefore, as in Section 1, we can express M as a product of powers of A_{ij}, B_{ij}, C_{ij} with exponents given by p, q, r and of an element of the commutator group of $\Gamma_g(2)$. Then the condition $\theta_{\mathfrak{m}}(M\cdot\tau)/\theta_{\mathfrak{n}}(M\cdot\tau) = \theta_{\mathfrak{m}}(\tau)/\theta_{\mathfrak{n}}(\tau)$ can be translated by Theorem 3 into

$$\sum_{i,j} p_{ij}m_i'm_j'' + \sum_{i\leqq j} q_{ij}m_i'm_j' + \sum_{i<j} r_{ij}m_i''m_j'' \equiv 0 \bmod 2$$

$$\sum_{i} q_{ii}(m_i')^2 + \sum_{i} r_{ii}(m_i'')^2 \equiv 0 \bmod 4$$

for all even $\mathfrak{m}$. It is an elementary exercise to see that this simply means $(q)_0\equiv(r)_0\equiv 0 \bmod 4$, $q\equiv r\equiv 0 \bmod 2$ and $p\equiv p_{11}\cdot 1_g \bmod 2$. Therefore either M or $-M$ is contained in $\Gamma_g(4,8)$, i.e. M is in $\Gamma_g(4,8)\,(\pm 1_{2g})$. This implies $\boldsymbol{C}(\theta_{\mathfrak{m}}/\theta_{\mathfrak{n}}) = F(\Gamma_g(4,8))$. Since $\boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}]$ contains elements of degree one, therefore, the field of fractions of $A(\Gamma_g(4,8))$ coincides with the field of fractions of $\boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}]$. This completes the proof of Theorem 5.

In order to derive some consequences from Theorem 5, we observe the following. Let R be an integral domain and suppose that a group Γ operates on R as a finite group of automorpisms. Then the invariant subring of the normalization of R is the normalization of the invariant subring of R (in the respective fields of fractions). Therefore, if Γ is a subgroup of $\Gamma_g(1)$ containing $\Gamma_g(4,8)$, then $A(\Gamma)$ is the normalization of the invariant subring of $\boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}]$ with respect to Γ. In case Γ is contained in $\Gamma_g(2)$, the invariant subring can be obtained in the following way. Suppose that $\psi_1,\cdots,\psi_n$ form a maximal set of linearly independent monomials in $\theta_{\mathfrak{m}}$ of degree $2k$, say, over $\boldsymbol{C}$. Then, for every M in $\Gamma_g(2)$, we have

$$M\cdot\Big(\sum_{1\leqq i\leqq n} c_i\psi_i\Big) = \sum_{1\leqq i\leqq n} c_iM\cdot\psi_i = \sum_{1\leqq i\leqq n} c_i\,\text{const.}\,\psi_i$$

where $c_1,\cdots,c_n$ are in $\boldsymbol{C}$. Therefore the invariant subring of $\boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}]$ is generated by invariant monomials. Furthermore, if we put

$$\psi = \theta_{\mathfrak{m}_1}\cdots\theta_{\mathfrak{m}_{2k}},$$

the invariance condition can be translated by Theorem 3 into some condition on the coefficients of $\mathfrak{m}_1,\cdots,\mathfrak{m}_{2k}$. If we take $\Gamma_g(4)$ as Γ, for instance, the condition is

$$\sum_{\alpha} m_{\alpha i}' \equiv \sum_{\alpha} m_{\alpha i}'' \equiv 0 \bmod 2$$

in which α runs over $1,\cdots,2k$. If we take $\Gamma_g(2)$ as Γ, the condition is

$$\begin{aligned}
&\sum_\alpha (m_{\alpha i}')^2 \equiv \sum_\alpha (m_{\alpha i}'')^2 \equiv 0 && \mod 4\\
&\sum_\alpha m_{\alpha i}' m_{\alpha j}' \equiv \sum_\alpha m_{\alpha i}'' m_{\alpha j}'' \equiv 0 && \mod 2\\
&\sum_\alpha m_{\alpha i}' m_{\alpha j}'' \equiv k \ \ (i=j), \equiv 0 \ \ (i \neq j) && \mod 2.
\end{aligned}$$

Also, if we take $\Gamma_g(2,4)$ as Γ, the first congruence becomes mod 2 and it will be contained in the second congruence. In any case, if we normalize the coefficients of the characteristics to 0, 1 and consider the multiplicities of even characteristics as unknown (non-negative) integers, the condition can be transformed into a finite number of homogeneous linear Diophantine equations. Therefore the invariant monomials form a finitely generated monoid and we can determine the generators.

In the simplest case $g=1$, the ring $\boldsymbol{C}[\theta_m\theta_n]$ is normal. In fact, the Riemann theta-formula implies

$$(\theta_{00})^4 - (\theta_{01})^4 - (\theta_{10})^4 = 0$$

and, since the polynomial $(X_{00})^4 - (X_{01})^4 - (X_{10})^4$ is absolutely irreducible, we see that $\boldsymbol{C}[\theta_m] = \boldsymbol{C}[\theta_{00}, \theta_{01}, \theta_{10}]$ is (canonically) isomorphic to the quotient ring of $\boldsymbol{C}[X]$ by the principal prime ideal generated by $(X_{00})^4 - (X_{01})^4 - (X_{10})^4$. We can very easily derive from this fact that $\boldsymbol{C}[\theta_m]$, hence also $\mathbf{C}[\theta_m\theta_n]$ are normal. Therefore, following the general procedure we explained and using a standard notation,[3] we get

$$\begin{aligned}
&A(\Gamma_1(4,8)) = \boldsymbol{C}[\theta_m]^{(2)} = \boldsymbol{C}[\theta_{00}, \theta_{01}, \theta_{10}]^{(2)}\\
&A(\Gamma_1(4)) = \boldsymbol{C}[(\theta_m)^2], \qquad A(\Gamma_1(2,4)) = \mathbf{C}[(\theta_m)^2]^{(2)}\\
&A(\Gamma_1(2)) = \boldsymbol{C}[(\theta_m)^4] = \boldsymbol{C}[(\theta_{01})^4, (\theta_{10})^4].
\end{aligned}$$

On the other hand, we observe that $\Gamma_1(1)/\Gamma_1(2)$ is isomorphic to the symmetric group π_3 of permutations on three letters, which we can identify with the three even characteristics mod 2, and the representation on $A(\Gamma_1(2))_2$ is its unique (up to equivalence) irreducible representation of degree two. Since $A(\Gamma_1(2))$ is (canonically) isomorphic to the graded ring of polynomials in two letters (each having degree two) with coefficients in $\boldsymbol{C}$, we know the action of $\Gamma_1(1)/\Gamma_1(2)$ on $A(\Gamma_1(2))$ completely. In particular, we have

---

[3] If S is a graded ring composed of homogeneous parts S_k with k running over an additive monoid J and if d is a positive integer, then

$$S^{(d)} = \bigoplus_{k \in J} S_{dk}.$$

$$\sum_{0\leqq k<\infty} \dim_{\mathbf{C}} A(\Gamma_1(1))_k t^k = 1/(1-t^4)(1-t^6).$$

Also we see that $A(\Gamma_1(1))$ is generated over $\boldsymbol{C}$ by

$$\psi_4 = \tfrac{1}{2}\cdot((\theta_{00})^8 + (\theta_{01})^8 + (\theta_{10})^8)$$
$$\psi_6 = \tfrac{1}{2}\cdot((\theta_{00})^4 + (\theta_{01})^4)((\theta_{00})^4 + (\theta_{10})^4)((\theta_{01})^4 - (\theta_{10})^4).$$

Because the Siegel operator Φ maps both ψ_4 and ψ_6 to 1, they coincide (including the constant factors) with Eisenstein series of weights four and six, which we shall explain in the Appendix. Moreover the kernel of Φ is the principal ideal generated by

$$(\psi_4)^3 - (\psi_6)^2 = 3^3/2^2\cdot(\theta_{00}\theta_{01}\theta_{10})^8.$$

The corresponding situation is fully investigated in the case $g=2$ in the paper mentioned in the introduction.

Appendix.

In this paper we gave a foundation for the construction of modular forms by theta-constants. Historically, this was the first method by which elliptic modular functions were (discovered and) investigated by Gauss. However, in the case of higher genus, the method originated by Eisenstein (and Poincaré) has been widely used. The essential point of this method can be stated as follows. Suppose that Δ is the subgroup of $Sp(g, \boldsymbol{R})$ consisting of elements with left bottom squares equal to zero. The group Δ is the stabilizer of the 0-dimensional boundary component $\lim i\eta 1_g$ for $\eta\rightarrow\infty$. At any rate, for any subgroup Γ of $Sp(g, \boldsymbol{R})$ commensurable with $\Gamma_g(1)$, we define an *Eisenstein series* belonging to Γ after Siegel by

$$\psi_k{}^{\Gamma}(\tau) = \sum \det(c\tau + d)^{-k}$$

in which k is an even integer and the summation is extended over elements M of Γ composed of a, b, c, d modulo left multiplications of elements of Δ. According to H. Braun, this series converges absolutely and uniformly in every compact subset of $\mathfrak{S}_g$ when and only when $k > g+1$. Moreover, for such values of k, the series defines a modular form of weight k belonging to Γ. Furthermore, as it follows from the work of Siegel [9], the ring $A(\Gamma)^{(2)}$ is the normalization of the graded ring generated over $\boldsymbol{C}$ by the Eisenstein series at least for $\Gamma = \Gamma_g(1)$. In the two special cases $g=1, 2$, we know also that $A(\Gamma_g(1))^{(2)}$ coincides with this graded ring of Eisenstein series. Now we shall show that this is no longer true for $g>2$. We need the following theorem:

THEOREM 6. *The group* $\Gamma_g(1)$ *is generated by* $\Delta \cap \Gamma_g(1)$ *and by the* g *matrices* D_i *equal to* E *for* $g = 1$ *in the* (i,i)-, $(i, i+g)$-, $(i+g, i)$-, $(i+g, i+g)$-*coefficients and to* 1_{2g} *in the remaining coefficients. Moreover, if*

$$M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

is in $\Delta \cap \Gamma_g(1)$, *we have* $\varkappa(M)^2 = \det(d) = \pm 1$ *and for* $i = 1, \cdots, g$ *we have* $\varkappa(D_i)^2 = e(\frac{1}{4})$.

The first part of the theorem is well known [cf. 7, p. 153] and is easy to prove. In the second part, the fact $\varkappa(M)^2 \det(d) = 1$ is straightforward while $\varkappa(D_i)^2 = e(\frac{1}{4})$ follows from Lemma 2.

Now, consider the following polynomial

$$\prod_{\mathfrak{m}\ \mathrm{even}} (t - (\theta_{\mathfrak{m}})^8) = t^{2^{g-1}(2^g+1)} + \cdots$$

in the indeterminate t. Then, using Theorem 6 and the transformation law of theta-functions, we see that the coefficients of this polynomial are invariant by the operations of $\Gamma_g(1)$. Hence they are modular forms belonging to $\Gamma_g(1)$. If we apply the Siegel operator g-times to the coefficients, the polynomial will become $(t-1)^{2^g}$. In particular

$$\psi = 1/2^g \cdot \sum_{\mathfrak{m} \bmod 2} (\theta_{\mathfrak{m}})^8$$

is a modular form of weight four belonging to $\Gamma_g(1)$ and it is mapped to 1 by the g-th power of the Siegel operator. Hence ψ is different from zero, and it is not contained in the graded ring of Eisenstein series for $g > 2$. This proves the assertion. Incidentally, the Riemann theta-formula shows that we have

$$\psi = 1/2^{g-1}(2^g - 1) \cdot \sum{}' e(\mathfrak{m}, \mathfrak{n}) (\theta_{\mathfrak{m}} \theta_{\mathfrak{n}})^4$$

where the summation is extended over all distinct pairs $(\mathfrak{m}, \mathfrak{n})$ of even characteristics mod 2. At any rate, the fact that Eisenstein series do not exist for smaller weights is a drawback and, although the method of Eisenstein-Poincaré is convenient to prove general qualitative theorems, the Gauss method to use theta-functions seems more suitable to investigate finer properties of modular varieties.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] Séminaire H. Cartan, Fonctions automorphes, E. N. S. (1957-58).

[2] G. Frobenius, "Ueber das Additionstheorem der Thetafunctionen mehrerer Variabeln," *Journal für die reine und angewandte Mathematik*, vol. 89 (1880), pp. 185-220.

[3] D. Hilbert, "Über die vollen Invariantensysteme," *Mathematische Annalen*, vol. 42 (1893), pp. 313-373.

[4] J. Igusa, "On Siegel modular forms of genus two," *American Journal of Mathematics*, vol. 84 (1962), pp. 175-200.

[5] ———, "Structure theorems of modular varieties," *Proceedings of the International Congress of Mathematicians* (1962), pp. 522-525.

[6] M. Koecher, "Zur Theorie de Modulfunktionen n-ten Grades, I," *Mathematische Zeitschrift*, vol. 59 (1954), pp. 399-416.

[7] A. Krazer, *Lehrbuch der Thetafunktionen*, Leipzig, 1903.

[8] A. Krazer-W. Wirtinger, "Abelsche Funktionen und algemeine Thetafunktionen," *Enzykl. Math. Wiss.*, II B 7, pp. 604-873.

[9] C. L. Siegel, "Einführung in die Theorie der Modulfunktionen n-ten Grades," *Mathematische Annalen*, vol. 116 (1939), pp. 617-657.

[10] ———, "On the theory of indefinite quadratic forms," *Annals of Mathematics*, vol. 45 (1944), pp. 577-622.

[11] A. Weil, *Foundations of algebraic geometry*, New York (1946).

[12] W. Wirtinger, *Untersuchungen über Thetafunktionen*, Leipzig (1895).

THE BASIC ESTIMATE OF THE $\bar\partial$-NEUMANN PROBLEM IN THE NON-KÄHLERIAN CASE.

By M. E. Ash.

Introduction. Let M be an open submanifold of a complex analytic manifold M' of complex dimension n. We assume that the closure of M is compact in M' and that the boundary bM of M is a $(2n-1)$-dimensional submanifold of M' of class C^∞. Denote by $\mathcal{A}^{p,q}$ the linear space of complex valued differential forms of type (p,q) on M, and let $\dot{\mathcal{A}}^{p,q}$ consist of those elements of $\mathcal{A}^{p,q}$ which are differentiable (C^∞) on the closure of M. If $\phi\in\mathcal{A}^{p,q}$, we note that the complex conjugate $\bar\phi$ of ϕ belongs to $\mathcal{A}^{q,p}$.

A Hermitian metric

$$ds^2 = h_{j\bar k}\,dz^j \otimes d\bar z^k \tag{1}$$

on M' allows us to define the Hermitian inner product $\phi\cdot\psi$ of $\phi,\psi\in\mathcal{A}^{p,q}$ at each point of M. There is then a map

$$*\colon \mathcal{A}^{p,q}\to\mathcal{A}^{n-q,n-p}$$

such that $*(1)$ is the volume element on M. The (global) Hermitian inner product and norm on $\mathcal{A}^{p,q}$ are defined by the equations

$$\begin{cases} (\phi,\psi)=\displaystyle\int_M \phi\cdot\psi*(1)=\int_M\phi\wedge*\bar\psi,\\ \|\phi\|^2=(\phi,\phi).\end{cases}$$

We have the usual operators

$$\begin{cases} \bar\partial\colon \mathcal{A}^{p,q}\to\mathcal{A}^{p,q+1}\\ \vartheta\colon \mathcal{A}^{p,q}\to\mathcal{A}^{p,q-1}\end{cases}$$

and the Green's formula

$$(\bar\partial\psi,\phi)-(\psi,\vartheta\phi)=\int_{bM}\psi\wedge*\bar\phi, \tag{2}$$
$$\psi\in\dot{\mathcal{A}}^{p,q-1},\ \phi\in\dot{\mathcal{A}}^{p,q}.$$

Received December 10, 1962.

Written while the author was a National Science Foundation Graduate Fellow. The author is indebted to Professors J. J. Kohn and D. C. Spencer for several helpful discussions concerning this paper.

Let $\phi \in \dot{\mathfrak{A}}^{p,q}$. The condition that $(\bar{\partial}\psi, \phi) = (\psi, \vartheta\phi)$ for all $\psi \in \dot{\mathfrak{A}}^{p,q-1}$ is, by this Green's formula, a boundary condition for ϕ.

In order to solve the $\bar{\partial}$-Neumann problem, J. J. Kohn [2, 3] has derived an *a priori* inequality for $\phi \in \dot{\mathfrak{A}}^{p,q}$ satisfying this boundary condition. The idea for such an inequality came from a paper of C. B. Morrey [4]. Kohn requires that the submanifold M of M' be strongly pseudoconvex and that the Herimitian metric be Kähler near bM. (Strong pseudoconvexity implies that such a metric exists.) The purpose of this note is to prove the inequality without this Kähler assumption (see § 4). Our method is to make the necessary computations in an orthonormal frame. In this setting certain difficulties involving the raising of indices are not encountered.

It might be mentioned that an analogous inequality was proved (following Kohn's technique) for complex multifoliate structure in the author's thesis [1].

1. **Orthonormal frames.** Let $(z^1, \cdots, z^n)$ be a holomorphic coordinate system on a coordinate neighborhood U. The differential forms $(dz^1, \cdots, dz^n)$ span the space of differential forms of type $(1,0)$. The inner product defined by a Hermitian metric (1) is given by $dz^j \cdot dz^k = h^{\bar{k}j}$, where $(h^{\bar{k}j})$ is the inverse of the matrix $(h_{l\bar{m}})$. We can apply the Gram-Schmidt orthonormalization process to obtain a local basis $(\omega^1, \cdots, \omega^n)$ of the space of differential forms of type $(1,0)$ such that $\omega^j \cdot \omega^k = \delta^{jk}$. We have $\omega^l = a_k{}^l dz^k$, where the $a_k{}^l$ are C^∞ functions.

In such an orthonormal frame the volume element has the expression

$$*(1) = (\sqrt{-1})^n \omega^1 \wedge \cdots \wedge \omega^n \wedge \bar{\omega}^1 \wedge \cdots \wedge \bar{\omega}^n.$$

A differential form ϕ of type (p,q) can be expressed in a moving frame as

$$\begin{aligned}\phi &= \phi_{J\bar{K}}\omega^J \wedge \bar{\omega}^K \\ &= \sum_{j_1 < \cdots < j_p} \sum_{k_1 < \cdots < k_q} \phi_{j_1 \cdots j_p \bar{k}_1 \cdots \bar{k}_q} \omega^{j_1} \wedge \cdots \wedge \omega^{j_p} \wedge \bar{\omega}^{k_1} \wedge \cdots \wedge \bar{\omega}^{k_q}.\end{aligned}$$

Here, J denotes an ordered set of integers $j_1 < \cdots < j_p$, and $\phi_{j_1 \cdots j_p \bar{k}_1 \cdots \bar{k}_q}$ is assumed to be antisymmetric in $j_1 \cdots j_p$ and $\bar{k}_1 \cdots \bar{k}_q$. The inner product at a point of $\phi, \psi \in \mathfrak{A}^{p,q}$ is given by

$$\phi \cdot \psi = \sum_{JK} \phi_{J\bar{K}} \bar{\psi}_{J\bar{K}}.$$

The element $*\phi$ has components

$$(*\phi)_{L\bar{M}} = (-1)^{np}(\sqrt{-1})^n \sum_{JK} \epsilon_{KL} \epsilon_{JM} \phi_{J\bar{K}}.$$

Here, $\epsilon_{i_1\cdots i_n}$ is the sign of the permutation $(i_1\cdots i_n)$ or zero if there are repeated indices.

Let $\frac{\partial}{\partial\omega^l}$, $\frac{\partial}{\partial\bar\omega^l}$ be the tangent vectors dual to ω^l, $\bar\omega^l$. This notation is not quite correct (see (4) below), but it is suggestive with regard to the calculations we shall be making. The operator $\bar\partial$ has the expression

$$\bar\partial\phi = \frac{\partial\phi_{JK}}{\partial\bar\omega^l}\bar\omega^l \wedge \omega^J \wedge \bar\omega^K + \cdots, \tag{3}$$

where the dots denote terms which do not involve the derivatives of the components of ϕ. These terms arise because, for instance,

$$\bar\partial\omega^k = c_{l\bar m}{}^k\omega^l \wedge \bar\omega^m$$

is not necessarily zero. We also have

$$\partial\bar\omega^k = -\bar c_{l\bar m}{}^k\omega^m \wedge \bar\omega^l.$$

Since $\bar\partial\partial g = -\partial\bar\partial g$ for any function g, we see that

$$\frac{\partial^2 g}{\partial\omega^l\partial\bar\omega^m} = \frac{\partial^2 g}{\partial\bar\omega^m\partial\omega^l} - \frac{\partial g}{\partial\omega^k}c_{l\bar m}{}^k + \frac{\partial g}{\partial\bar\omega^k}\bar c_{ml}{}^k. \tag{4}$$

The operator ϑ has the expression

$$\vartheta\phi = (-1)^{p+1}\sum_l \frac{\partial\phi_{Jl\bar K}}{\partial\omega^l}\,\omega^J \wedge \bar\omega^K + \cdots \tag{5}$$

We have shown how the calculus of differential forms on a complex analytic manifold appears in an orthonormal frame. Conversely an almost complex manifold is essentially a C^∞ manifold which is covered by such frames on which the operators ∂ and $\bar\partial$ are defined. The integrability condition is $\bar\partial^2 = 0$, or, equivalently, $\bar\partial\partial = -\partial\bar\partial$. (See [5], pp. 30-35.) Therefore, in the following we shall asume that M' is an integrable almost complex manifold with a Hermitian metric.

2. Strong pseudoconvexity. Let $|r(x)|$ be the geodesic distance of $x \in M'$ to bM, and suppose that $r(x)$ is negative if $x \in M$ and positive if $x \in M' - \bar M$. The volume element on bM is $*(dr)$.

Definition. The submanifold M of M' is strongly pseudoconvex if and only if for any tangent vector τ of type $(1,0)$ at a point of bM, the condition $\langle\partial r, \tau\rangle = 0$ implies that $\langle\partial\bar\partial r, \tau \wedge \bar\tau\rangle > 0$.

The symbol $\langle\ ,\ \rangle$ denotes the contraction of tangent vectors and differential forms. Note that

$$\langle \partial\bar{\partial} r, \tau \wedge \bar{\tau}\rangle = \langle \bar{\partial}\partial r, \bar{\tau} \wedge \tau\rangle$$

on an integrable almost complex manifold. The function r in the definition could actually be any function such that $r < 0$ on M, $r = 0$ on bM, and $r > 0$ on $M' - M$, with $dr \neq 0$ on bM.

It is a simple computation to show that the strong pseudoconvexity condition above can be expressed in a frame in the following manner: at each point of bM the condition $\tau^l \dfrac{\partial r}{\partial \omega^l} = 0$ implies

$$\left(\frac{\partial^2 r}{\partial \omega^l \partial \bar{\omega}^m} - \bar{c}_{m\bar{l}}^{k} \frac{\partial r}{\partial \bar{\omega}^k}\right) \tau^l \bar{\tau}^m > 0. \tag{6}$$

3. Integration by parts. Let f and g be functions on M', and suppose that they have support in a coordinate neighborhood so that our manipulations in a given orthonormal frame have meaning. Then

$$\begin{aligned} d[f\bar{g}(\sqrt{-1})^n \omega^1 \wedge \cdots \wedge \omega^{l-1} \wedge \omega^{l+1} \wedge \cdots \wedge \omega^n \wedge \bar{\omega}^1 \wedge \cdots \wedge \bar{\omega}^n] \\ = (-1)^{l-1} \frac{\partial f}{\partial \omega^l} \bar{g} * (1) + (-1)^{l-1} f \frac{\partial \bar{g}}{\partial \omega^l} * (1) \\ + (-1)^{l-1} f \bar{g} S_l * (1), \end{aligned}$$

where S_l is defined by this equation. Let T denote the tangential part of a differential form on bM. Then

$$\begin{aligned} T[(\sqrt{-1})^n \omega^1 \wedge \cdots \wedge \omega^{l-1} \wedge \omega^{l+1} \wedge \cdots \wedge \omega^n \wedge \bar{\omega}^1 \wedge \cdots \wedge \bar{\omega}^n] \\ = \lambda * (dr). \end{aligned}$$

Since

$$dr \wedge *(dr) = dr \cdot dr * (1) = *(1),$$

we see that

$$\lambda = (-1)^{l-1} \frac{\partial r}{\partial \omega^l}.$$

From the above equations and Stokes' theorem we obtain

$$\begin{aligned} \int_M \frac{\partial f}{\partial \omega^l} \bar{g} * (1) = -\int_M f \frac{\partial \bar{g}}{\partial \omega^l} * (1) + \int_{bM} f \bar{g} \frac{\partial r}{\partial \omega^l} * (dr) \\ - \int_M f \bar{g} S_l * (1). \end{aligned} \tag{7}$$

Similarly, we have (integrations are over M unless otherwise specified)

$$(8)\qquad \int f\,\frac{\partial\bar g}{\partial\bar\omega^m}*(1)=-\int\frac{\partial f}{\partial\bar\omega^m}\,\bar g*(1)+\int_{bM}f\bar g\,\frac{\partial r}{\partial\bar\omega^m}*(dr)-\int f\bar g S_{\bar m}*(1)$$

Applying (7), then (4), and finally (8), we obtain the following formula for integration by parts

$$(9)\qquad \begin{aligned}\int\frac{\partial f}{\partial\omega^l}\,\frac{\partial\bar g}{\partial\bar\omega^m}*(1)=\int\frac{\partial f}{\partial\bar\omega^m}\,\frac{\partial\bar g}{\partial\omega^l}*(1)\\ +\int_{bM}f[\frac{\partial\bar g}{\partial\bar\omega^m}\,\frac{\partial r}{\partial\omega^l}-\frac{\partial\bar g}{\partial\omega^l}\,\frac{\partial r}{\partial\bar\omega^m}]*(dr)\\ -\int_{bM}f\bar g[\bar c_{ml}^{k}\,\frac{\partial r}{\partial\bar\omega^k}+S_l\,\frac{\partial r}{\partial\bar\omega^m}]*(dr)+W,\end{aligned}$$

where

$$\begin{aligned}W=\int f\,\frac{\partial\bar g}{\partial\omega^l}S_{\bar m}*(1)+\int f\,\frac{\partial\bar g}{\partial\omega^k}c_{l\bar m}^{k}*(1)+\int\frac{\partial f}{\partial\bar\omega^k}\,\bar g\bar c_{ml}^{k}*(1)\\ +\int f\bar g\,\frac{\partial\bar c_{ml}^{k}}{\partial\bar\omega^k}*(1)+\int f\bar g\bar c_{ml}^{k}S_{\bar k}*(1)+\int\frac{\partial f}{\partial\bar\omega^m}\,\bar g S_l*(1)\\ +\int fg\,\frac{\partial S_l}{\partial\bar\omega^m}*(1)+\int f\bar g S_l S_{\bar m}*(1).\end{aligned}$$

Note that for any functions σ, τ and positive number C we have

$$|\int\sigma\tau|\leqq\int|\frac{1}{C}\sigma|\,|C\tau|\leqq\frac{1}{2C^2}\int|\sigma|^2+\frac{C^2}{2}\int|\tau|^2.$$

Thus, we see that for any positive number ϵ there is a positive number $K(\epsilon)$ such that

$$(10)\qquad |W|\leqq\epsilon\sum_k\int[|\frac{\partial f}{\partial\bar\omega^k}|^2+|\frac{\partial g}{\partial\bar\omega^k}|^2]*(1)+K(\epsilon)\int[|f|^2+|g|^2]*(1).$$

4. The basic $\bar\partial$-estimate. The considerations used in § 3 of [3] to get the boundary condition from Green's formula (2) are exactly the same in an orthonormal frame. Thus $\phi\in\dot{\mathfrak{A}}^{p,q}$ satisfies the condition

$$(\bar\partial\psi,\phi)=(\psi,\vartheta\phi)$$

for all $\psi\in\dot{\mathfrak{A}}^{p,q-1}$ if and only if at each point of bM

$$(11)\qquad \sum\phi_{J\bar m\bar K}\,\frac{\partial r}{\partial\omega^m}=0.$$

Let U be a coordinated neighborhood in M' and consider $\phi\in\dot{\mathfrak{A}}^{p,q}$ restricted to U. We define the absolute value of ϕ at each point of U by the formula

$$|\phi|^2 = \sum_{J\bar K} |\phi_{J\bar K}|^2$$

and the norm of all barred derivatives of ϕ by the equation

$$\|\phi\|_{\bar\omega}^2 = \sum_{JKl} \int_{M\cap U} \left|\frac{\partial\phi_{J\bar K}}{\partial\bar\omega^l}\right|^2 *(1).$$

Then we have the following result (see [3], Theorem 5.6):

THEOREM. *Let M' be an integrable almost complex manifold with a Hermitian metric and suppose that M is a strongly pseudoconvex submanifold of M' with a regular boundary bM. For each point of bM there is a coordinate neighborhood U in M' and a positive number C such that if $\phi \in \dot{\mathfrak{A}}^{p,q}$, $q > 0$, has support in U and satisfies boundary condition* (11), *then*

$$\|\bar\partial\phi\|^2 + \|\vartheta\phi\|^2 + \|\phi\|^2 \geqq C(\|\phi\|_{\bar\omega}^2 + \int_{bM} |\phi|^2 *(dr) + \|\phi\|^2).$$

Proof. If $\sigma, \tau \in \dot{\mathfrak{A}}$ and C is a positive number, then

$$0 \leqq \left\|\frac{1}{C}\sigma + C\tau\right\|^2 = \frac{1}{C^2}\|\sigma\|^2 + C^2\|\tau\|^2 + 2\operatorname{Re}(\sigma,\tau).$$

Letting $C^2 = 2$, we see that

$$\begin{aligned}\|\sigma+\tau\|^2 &= \|\sigma\|^2 + \|\tau\|^2 + 2\operatorname{Re}(\sigma,\tau)\\ &\geqq \tfrac{1}{2}\|\sigma\|^2 - \|\tau\|^2.\end{aligned} \tag{12}$$

Thus, (5) gives that there is a positive number A such that

$$\|\vartheta\phi\|^2 \geqq \tfrac{1}{2} \sum_{JKlm} \int \frac{\partial\phi_{J\bar l\bar K}}{\partial\omega^l} \frac{\partial\bar\phi_{J\bar m\bar K}}{\partial\bar\omega^m} *(1) - A\|\phi\|^2.$$

Applying formula (9) for integration by parts, we obtain

$$\begin{aligned}\|\vartheta\phi\|^2 \geqq{} & \tfrac{1}{2}\sum \int \frac{\partial\phi_{J\bar l\bar K}}{\partial\bar\omega^m} \frac{\partial\bar\phi_{J\bar m\bar K}}{\partial\omega^l} *(1)\\ & + \tfrac{1}{2}\sum \int_{bM} \phi_{J\bar l\bar K}\left[\frac{\partial\bar\phi_{J\bar m\bar K}}{\partial\bar\omega^m} \frac{\partial r}{\partial\omega^l} - \frac{\partial\bar\phi_{J\bar m\bar K}}{\partial\omega^l} \frac{\partial r}{\partial\bar\omega^m}\right] *(dr)\\ & - \tfrac{1}{2}\sum \int_{bM} \phi_{J\bar l\bar K}\bar\phi_{J\bar m\bar K}\left[\bar c_{ml}{}^k \frac{\partial r}{\partial\bar\omega^k} + S_l \frac{\partial r}{\partial\bar\omega^m}\right] *(dr) + W.\end{aligned} \tag{13}$$

where, by (10), for any positive number ϵ there is a positive number $K(\epsilon)$ such that

$$|W| \leqq \epsilon\|\phi\|_{\bar\omega}^2 + K(\epsilon)\|\phi\|^2. \tag{14}$$

Note that we have absorbed the term $A \,\| \phi \|^2$ into W.

Differentiating the conjugate of boundary condition (11) with respect to ω^l, we have

$$\sum_m \frac{\partial \bar{\phi}_{J\bar{m}\bar{K}}}{\partial \omega^l} \frac{\partial r}{\partial \bar{\omega}^m} + \sum_m \bar{\phi}_{J\bar{m}\bar{K}} \frac{\partial^2 r}{\partial \omega^l \partial \bar{\omega}^m} = \lambda \frac{\partial r}{\partial \omega^l}.$$

This is because if $f = 0$ on bM, then $df = \lambda\, dr$ on bM. Multiplying by $\phi_{J\bar{l}\bar{K}}$ and summing on l, we see that

$$\sum_{lm} \phi_{J\bar{l}\bar{K}} \frac{\partial \bar{\phi}_{J\bar{m}\bar{K}}}{\partial \omega^l} \frac{\partial r}{\partial \bar{\omega}^m} + \sum_{lm} \phi_{J\bar{l}\bar{K}} \bar{\phi}_{J\bar{m}\bar{K}} \frac{\partial^2 r}{\partial \omega^l \partial \bar{\omega}^m}$$
$$= \lambda \sum_l \phi_{J\bar{l}\bar{K}} \frac{\partial r}{\partial \omega^l} = 0.$$

This last step follows from (11). Thus we have

$$-\sum_{lm} \phi_{J\bar{l}\bar{K}} \Big[\frac{\partial \bar{\phi}_{J\bar{m}\bar{K}}}{\partial \omega^l} \frac{\partial r}{\partial \bar{\omega}^m} + \bar{\phi}_{J\bar{m}\bar{K}} \bar{c}_{ml}{}^k \frac{\partial r}{\partial \bar{\omega}^k} \Big] = \sum_{lm} F_{l\bar{m}} \phi_{J\bar{l}\bar{K}} \bar{\phi}_{J\bar{m}\bar{K}},$$

where

$$F_{l\bar{m}} = \frac{\partial^2 r}{\partial \omega^l \partial \bar{\omega}^m} - \bar{c}_{ml}{}^k \frac{\partial r}{\partial \bar{\omega}^k}.$$

Boundary condition (11) and strong pseudoconvexity condition (6) give

$$\sum_{lm} F_{l\bar{m}} \phi_{J\bar{l}\bar{K}} \bar{\phi}_{J\bar{m}\bar{K}} > 0.$$

Since $(F_{l\bar{m}})$ is a Hermitian matrix (see (4)), there is a positive number C such that

$$-\sum_{lm} \phi_{J\bar{l}\bar{K}} \Big[\frac{\partial \bar{\phi}_{J\bar{m}\bar{K}}}{\partial \omega^l} \frac{\partial r}{\partial \bar{\omega}^m} + \bar{\phi}_{J\bar{m}\bar{K}} \bar{c}_{ml}{}^k \frac{\partial r}{\partial \bar{\omega}^k} \Big] \tag{15}$$
$$\geqq 2C \sum_l | \phi_{J\bar{l}\bar{K}} |^2.$$

Boundary condition (11) shows that the first and fourth boundary integrals in (13) are zero, and hence we obtain from (15) that

$$\| \vartheta\phi \|^2 \geqq \tfrac{1}{2} \sum \int \frac{\partial \phi_{J\bar{l}\bar{K}}}{\partial \bar{\omega}^m} \frac{\partial \bar{\phi}_{J\bar{m}\bar{K}}}{\partial \omega^l} *(1) \tag{16}$$
$$+ C \int_{bM} | \phi |^2 *(dr) + W.$$

From (3) and (12) we see that there is a positive number A such that

$$
\begin{aligned}
\| \bar{\partial}\phi \|^2 &\geqq \tfrac{1}{2} \left\| \frac{\partial \phi_{J\bar{L}}}{\partial \overline{\omega}^m} \, \overline{\omega}^m \wedge \omega^J \wedge \overline{\omega}^L \right\|^2 - A \| \phi \|^2 \\
&= \tfrac{1}{2} \sum_{JLm} \int \left| \frac{\partial \phi_{J\bar{L}}}{\partial \overline{\omega}^m} \right|^2 *(1) \\
&\quad + \tfrac{1}{2} \sum_J \sum_{(mL) \neq (lM)} \int \frac{\partial \phi_{J\bar{L}}}{\partial \overline{\omega}^m} \, \frac{\partial \bar{\phi}_{J\bar{M}}}{\partial \omega^l} \, \epsilon_{lM}{}^{mL} *(1) - A \| \phi \|^2.
\end{aligned}
$$

Here, (mL) represents the unordered sequence $ml_1 \cdots l_q$. Let $\langle mL \rangle$ denote the set (mL) ordered in increasing order. Then

$$
\epsilon_{lM}{}^{mL} = \begin{cases} 0 & \text{if } \langle mL \rangle \neq \langle lM \rangle \\ -\epsilon^{\,lK}_{\langle lK \rangle} \; \epsilon^{\,mK}_{\langle mK \rangle} & \text{if } L = \langle lK \rangle,\ M = \langle mK \rangle \end{cases}
$$

and we have

$$
\begin{aligned}
\| \bar{\partial}\phi \|^2 \geqq \tfrac{1}{2} \| \phi \|_{\bar{\omega}}^2 - \tfrac{1}{2} \sum \int \frac{\partial \phi_{J l \bar{K}}}{\partial \overline{\omega}^m} \, \frac{\partial \bar{\phi}_{J m \bar{K}}}{\partial \omega^l} \, *(1) \\
- A \| \phi \|^2.
\end{aligned} \tag{17}
$$

Inequalities (16) and (17) combine to give

$$
\begin{aligned}
\| \bar{\partial}\phi \|^2 + \| \vartheta\phi \|^2 \geqq \tfrac{1}{2} \| \phi \|_{\bar{\omega}}^2 + C \int_{bM} | \phi |^2 *(dr) \\
+ W - A \| \phi \|^2.
\end{aligned}
$$

By (14), this completes the proof of the Theorem.

PRINCETON UNIVERSITY.
BRANDEIS UNIVERSITY.

REFERENCES.

[1] M. E. Ash, "The Neumann problem on strongly pesudo-convex complex multifoliate manifolds," Thesis, Princeton (December 1962).

[2] J. J. Kohn, "Solution of the $\bar{\partial}$-Neumann problem on strongly pseudo-convex manifolds," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 47 (1961), pp. 1198-1202.

[3] ———, "Harmonic integrals on strongly pseudo-convex manifolds, I," *Annals of Mathematics*, vol. 78 (1963), pp. 112-148.

[4] C. B. Morrey, "The analytic imbedding of abstract real-analytic manifolds," *Annals of Mathematics*, vol. 68 (1958), pp. 159-201.

[5] A. Weil, *Introduction à l'étude des variétés kählériennes* (1958).

ZEROES OF POLYNOMIALS OVER FINITE FIELDS.*

By James Ax.

1. Introduction.

Let $F = F(X_1, \cdots, X_n)$ be a polynomial of (total) degree d over a finite field k with q elements. In section **3**, making use of some ideas of B. Dwork in [2], we prove the following theorem:

If b is the largest integer (strictly) less than n/d then q^b divides the number of zeroes of F.

E. Artin had conjectured that if F is homogeneous and $n > d$ then F has a non-trivial zero. C. Chevalley proved this in [1] and even showed the hypothesis of homogeneity could be replaced by the weaker assumption of no constant term. E. Warning in [4], using a lemma of Chevalley, showed that even without this last assumption the characteristic p of k divides $N(F)$, the number of zeroes of F (counting the trivial zero if F has no constant term). In Section **2** we give a quick proof of the Chevalley-Warning theorem independent of the Chevalley lemma. Nevertheless, there does not seem to be any simple proof of the fact that q divides $N(F)$ if $n > d$.

In Section **4** we exhibit, for each n and d, a polynomial of degree d in n variables such that the highest power of p dividing the number of its zeroes is precisely q^b if b is the largest integer less than n/d. While our result is the best possible divisibility relation in this sense, E. Warning in [4] showed that if $n > d$ and if F has at least one zero then $N(F)$ is at least q^{n-d}.

The zeta function $Z(H;t)$ of the hypersurface H defined by F over k is defined by

$$Z(H;t) = \exp\left(\sum_{s=1}^{\infty} N_s t^s/s\right)$$

where N_s is the number of zeroes of F in the field with q^s elements. Let Ω denote the completion of the algebraic closure of the p-adic completion of the rationals, and let $|\ |$ be the valuation on Ω normed so that $|p| = 1/p$.

The referee has shown how our result may be reformulated as the following statement, using the above notation.

Received February 15, 1963.

Revised January 7, 1964.

* Written with the partial support of NSF Grant G23834.

THEOREM. *Each pole and each zero of* $Z(H;t)$ *has p-adic valuation at least* q^b.

Indeed, it follows from the rationality of $Z(H;t)$ [2], that

$$\exp\left(\sum_{s=1}^{\infty} N_s t^s/s\right) = \prod_i (1-\alpha_i t)/\prod_j (1-\beta_j t)$$

where the α_i, β_j are algebraic integers, $\alpha_i \neq \beta_j$, for i and j ranging over finite sets. By logarithmic differentiations we obtain

$$\sum_{s=1}^{\infty} N_s t^{s-1} = \sum_j \beta_j(1-\beta_j t)^{-1} - \sum_i \alpha_i(1-\alpha_i t)^{-1}.$$

If we now assume our result, $|N_s| \leqq |q^{bs}| = q^{-bs}$, so that the left side converges in Ω for $|t| < q^b$, then $|\alpha_i|$, $|\beta_j| \leqq q^{-b}$ which verifies the referee's statement. The converse follows from

$$N_1 = \sum_j \beta_j - \sum_i \alpha_i.$$

Throughout this paper, F, $N(F)$, n, d, b, k, q, and p are above. Z denotes the integers, Z_+ the nonnegative integers. If $u = (u_1, \cdots, u_r) \in (Z_+)^r$, X^u denotes the monomial $\prod_{i=1}^{r} X_i^{u_i}$ and we define height $u =$ degree $X^u = \sum_{i=1}^{r} u_i$.

2. Quick proof of the Chevalley-Warning theorem.

Since each element of k is a $q-1$ root of unity or zero, we have for each $x \in k^n$ that $1 - F(x)^{q-1} = 1$ if $F(x) = 0$, zero otherwise. Summing over $x \in k^n$, we have (in k)

$$(1)^{[3]} \qquad N(F) = \sum(1-F(x)^{q-1}) = -\sum F(x)^{q-1}.$$

Now F^{q-1}, being of degree $d(q-1)$ is a k-linear combination of monomials of degree at most $d(q-1)$. If X^u is such a monomial, we compute

$$(2) \qquad \sum_{x \in k^n} x^u = \prod_{i=1}^{n} \sum_{x_i \in k} x_i^{u_i} = \prod_{i=1}^{n} Y(u_i)$$

where $Y(u_i) = q-1 = -1$ if u_i is positive multiple of $q-1$, zero otherwise. If $d < n$, then height $u \leqq d(q-1) < n(q-1)$ which implies that the sum in (2) is zero. Hence, the sum in (1) is zero, i.e., $N(F) \equiv 0 \bmod p$.

3. Proof of the theorem.

Let $q = p^f$, Q_p be the p-adic completion of the rationals, and K the unique unramified extension of Q_p of degree f. Then the residue class field of K is k. Let T denote the set of Teichmüller representatives of k in K; let $T^* = T - \{0\}$, the $q-1$ roots of unity. Let ζ be a primitive p-th root of unity. If α is an integer of Q_p, ζ^α is defined to be ζ^a if $a(\in Z_+)$ is congruent to α modulo p. Letting S denote the trace of K over Q_p we define $C = \sum_{m=0}^{q-1} c(m) U^m$ to be the unique polynomial of degree $q-1$ with coefficients in $K(\zeta)$ such that $C(t) = \zeta^{S(t)}$ for all $t \in T$. Summing $C(t)t^{-j}$ over $t \in T^*$ we find

---

[3] This equation is the essential fact in our proof as in Warning's. We then proceed directly to the result in a way suggestive of certain manipulations in the sequel.

(3) $c(j)(q-1) = g(j)$ for $0 < j < q-1$ where the Gauss sum $g(j)$ is defined for $0 \leqq j < q-1$ by

$$g(j) = \sum t^{-j}\zeta^{S(t)} \qquad (t \in T^*).$$

Summing $C(t)$ over $t \in T^*$ and using that the trace function is not identically zero on a finite field, we find

$$-1 = g(0) = (q-1)(c(0) + c(q-1)).$$

Since

$$c(0) = 1, \tag{3'}$$

we have

$$c(q-1)(q-1) = -q. \tag{3''}$$

If $0 \leqq j \leqq q-1$, let j_i for $i = 0, \cdots, f-1$ be such that $0 \leqq j_i \leqq p-1$ and $j = \sum_{i=0}^{f-1} j_i p^i$. We set $\sigma(j) = \sum_{i=0}^{f-1} j_i$, $\rho(j) = \prod_{i=0}^{f-1} j_i!$ and $\lambda = \zeta - 1$. Then Stickelberger's congruence [3] (and [2] for further reference),

$$g(j)\rho(j)/\lambda^{\sigma(j)} \equiv -1 \mod \lambda \text{ for } 0 \leqq j < q-1$$

together with (3), (3′), and (3″) certainly imply

$$c(j) \equiv 0 \mod \lambda^{\sigma(j)} \text{ for } 0 \leqq j \leqq q-1. \tag{4}$$

The map $\alpha \to \zeta^{S(\alpha)}$ is a non-trivial character of the additive group of the integers of K, trivial on the maximal ideal of the integers of K. Thus the map β from k to the p-th roots of unity defined by $\beta(x) = C(t)$ for $x \in k$ and t the Teichmüller representative of x is a non-trivial character of the additive group of k. If $u \in k$, then $\sum \beta(x_0 u) = q$ if $u = 0$, zero otherwise where the sum if over $x_0 \in k$. It follows that

$$qN(F) = \sum \beta(x_0 F(x_1, \cdots, x_n)) \qquad ((x_0, \cdots, x_n) \in k^{n+1}).$$

Let

$$F=\sum a(w)X^{w}\qquad (w\in W)$$

where W is the set of $w\in(Z_+)^n$ such that height $w\leqq d$. We have

$$qN(F)=\sum\prod_{w\in W}\beta(a(w)x^{w'})\qquad (x=(x_0,\cdots,x_n)\in k^{n+1})$$

where if $w=(w_1,\cdots,w_n)\in(Z_+)^n$ then $w'=(1,w_1,\cdots,w_n)\in(Z_+)^{n+1}$. If $A(w)$ is the Teichmüller representative of $a(w)$ for each $w\in W$, then

$$qN(F)=\sum\prod_{w\in W}C(A(w)t^{w'})\qquad (t=(t_0,\cdots,t_n)\in T^{n+1})$$

and so

$$\begin{aligned}qN(F)&=\sum_{t\in T^{n+1}}\prod_{w\in W}\sum_{m=0}^{q-1}c(m)A(w)^{m}t^{mw'}\\ &=\sum_{m\in M}\sum_{t\in T^{n+1}}\prod_{w\in W}(c(m(w))A(w)^{m(w)}t^{mw'})\end{aligned}\tag{5}$$

where M is the set of functions on W with values from the integers $0,1,\cdots,q-1$. Setting $\alpha(m)=\prod_{w\in W}A(w)^{m(w)}\in T$, $e(m)'=\sum_{w\in W}m(w)w'$ for $m\in M$, we may rewrite (5) as

$$qN(F)=\sum_{m\in M}\alpha(m)\prod_{w\in W}c(m(w))\sum_{t\in T^{n+1}}t^{e(m)'}.\tag{5'}$$

If $v\in(Z_+)^r$ we write $q-1\mid v$ if there exists $u\in(Z_+)^r$ such that $v=(q-1)u$ and $q-1\nmid v$ otherwise. Let m be an arbitrary element of M. Then we easily compute

$$\sum_{t\in T^{n+1}}t^{e(m)'}=0 \text{ if } q-1\nmid e(m)'\tag{6}$$

and

$$\sum_{t\in T^{n+1}}t^{e(m)'}=q^{n+1} \text{ if } e(m)'=(0,\cdots,0).\tag{6'}$$

We now assume $q-1\mid e(m)'$ and $e(m)'\neq(0,\cdots,0)$, i.e., $m(w)\neq 0$ for some $w\in W$. Let $e(m)=\sum_{w\in W}m(w)w$ and let s be the number of non-zero entries in $e(m)$, $0\leqq s\leqq n$. We have

$$\sum_{t\in T^{n+1}}t^{e(m)'}=(q-1)^{s+1}q^{n-s} \text{ if } q-1\mid e(m)'\neq(0,\cdots,0)\tag{6''}$$

taking into account that the first entry of $e(m)'$, $\sum_{w\in W}m(w)$, is a non-zero multiple of $q-1$. For each $w\in W$, let $m_i(w)$ for $i=0,\cdots,f-1$ be such that $m(w)=\sum_{i=0}^{f-1}m_i(w)p^i$ and $0\leqq m_i(w)\leqq p-1$. We extend the definition of $m_z(w)$ to all $z\in Z$ by letting $m_z(w)=m_r(w)$ if r is the least non-negative residue of z modulo f and define for each $j=0,\cdots,f-1$ the function $m^{(j)} \in M$ by

$$m^{(j)}(w) = \sum_{i=0}^{f-1} m_{i-j}(w) p^i. \tag{7}$$

Using $t^q = t$ for all $t \in T$, we readily compute that the effect of substituting $m^{(j)}$ for m in the sum in (6″) is the same as if we formally substitute t^{p^j} for t, i.e., no change since $t \to t^p$ is a permutation of T. We deduce from the mutually exclusive (6), (6′) and (6″) that $q-1 \mid e(m^{(j)})'$, and the number of non-zero entries of $e(m^{(j)})$ is again s for each $j=0,\cdots,f-1$. This yields the inequalities

$$s(q-1) \leqq \text{height}\, e(m^{(j)}) = \text{height} \sum_{w \in W} m^j(w) w \leqq d \sum_{w \in W} m^{(j)}(w).$$

Since $\sum_{w \in W} m^{(j)}(w)$, the first entry of $e(m^{(j)})'$, is a multiple of $q-1$ we conclude

$$(s/d)^*(q-1) \leqq \sum_{w \in W} m^{(j)}(w),$$

where $(y)^*$ means the smallest integer not less than y. Summing this relation over $j=0,\cdots,f-1$, using (7) and interchanging order of summation twice we obtain

$$f(s/d)^*(q-1) \leqq \sum_{w \in W} \sum_{i=0}^{f-1} p^i \sum_{j=0}^{f-1} m_{i-j}(w).$$

Thus with σ as used in (4)

$$f(s/d)^*(q-1) \leqq \sum_{w \in W} \sum_{i=0}^{f-1} p^i \sigma(m(w)).$$

So

$$f(p-1)(s/d)^* \leqq \sum_{w \in W} \sigma(m(w))$$

which in view of (4) and the fact that p divides λ^{p-1} implies the exponent of the highest power of q dividing $\prod_{w \in W} c(m(w))$ is at least $(s/d)^*$. Combining this with (6), (6′), and (6″) we see from (5′) that

$$q^r \mid q(N(F) \quad \text{if} \quad r = \min r(s) \tag{8}$$

where

$$r(s) = (s/d)^* + n - s, \quad s = 0, 1, \cdots, n. \tag{9}$$

Now

$$h \geqq ((s+h)/d)^* - (s/d)^*, \quad h \in Z_+$$

since in going from h to $h+1$ the left side increases by one while the right side increases by at most one. Substituting $h = n - s$ in the relation and using $b+1=(n/d)^*$ we see from (9) that $r(s)\geqq b+1$ for $s=0,\cdots,n$. By (8) q^b divides $N(F)$.

COROLLARY. *If F_i is a polynomial in n variables of degree d_i for $i=1,\cdots,j$ then the number N of common zeroes of the F_i is divisible by q^b if $n>b\sum_{i=0}^{j}d_i$.*

Proof. A standard combinatorial argument shows

$$N=-\sum_{s}(-1)^{\#S}N(\prod_{i\in S}F_i)$$

where the sum is over all non-empty subsets S of the set of integers $1,\cdots,j$ and where $\#S=$ number of elements of S. The corollary follows from the theorem since for each S,

$$\text{degree}\prod_{i\in S}F_i\leqq\sum_{i=1}d_i.$$

4. Examples. If $a\in Z_+$, $a>0$ we define

$$G_{a,d}(X_1,\cdots,X_{ad})=X_1\cdots X_d+\cdots+X_{(a-1)d+1}\cdots X_{ad}$$

and assert that the highest power of p dividing $N(G_{a,d})$ is q^{a-1}. Now $N(G_{1,d})=q^d-(q-1)^d$. $N(G_{a+1,d})=N(G_{a,d})$ times the number of zeroes of $X_{(a-1)d+1}\cdots X_{ad}$ (in k^d) plus the number of non-zeroes of $G_{a,d}$ (in k^{ad}) times the (constant) number of representations of a non-zero element of k by $X_{(a-1)d+1}\cdots X_{ad}$ (in k^d), i.e.,

$$\begin{aligned}N(G_{a+1,d})&=N(G_{a,d})N(G_{1,d})+(q^{ad}-N(G_{a,d}))(q-1)^{d-1}\\&=qN(G_{a,d})(q^{d-1}-(q-1)^{d-1})+q^{ad}(q-1)^{d-1}.\end{aligned}\tag{10}$$

For $d>1$ this yields our assertion recursively; for $d=1$ our assertion is immediate. If $n=bd+h$ with $0<h\leqq d$ (so that b is largest integer less than n/d) we set

$$F(X_1,\cdots,X_n)=G_{b,d}(X_1,\cdots,X_{bd})\ \text{if}\ h=1,$$

$$F(X_1,\cdots,X_n)=G_{b,d}(X_1,\cdots,X_{bd})+X_{bd+1}\cdots X_n\ \text{if}\ h>1.$$

We assert that the highest power of p dividing $N(F)$ is q^b.

If $h=1$ this follows from our previous assertion since in this case $N(F)=qN(G_{b,d})$. If $h>1$ our previous assertion still yields the desired result since by reasoning similar to that used in establishing (10) we have

$$N(F)=qN(G_{b,d})(q^{h-1}-(q-1)^{h-1})+q^{bd}(q-1)^{h-1}.$$

REFERENCES.

[1] C. Chevalley, "Démonstration d'une hypothèse de M. Artin," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 11 (1936), pp. 73-75.

[2] B. Dwork, "On the rationality of the zeta function of an algebraic variety," *American Journal of Mathematics*, vol. 82 (1960), pp. 631-648.

[3] L. Stickelberger, "Über eine Verallgemeinerung der Kreistheilung," *Mathematische Annalen*, vol. 37 (1890), pp. 321-367.

[4] E. Warning, "Bermerkung zur vorstehenden Arbeit von Herrn Chevalley," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 11 (1936), pp. 76-83.

A NOTE ON HERMITIAN FORMS OVER FIELDS OF CHARACTERISTIC 2.

By Chih-Han Sah.*

In a short note, Jacobson [5] showed that over "nice" fields of characteristic $\neq 2$ with an involution, the problem of field equivalence between hermitian forms may be reduced to one of field equivalence between two quadratic forms. Using this, he obtained a complete set of field invariants for hermitian forms over a large family of fields. Recently, Jacobowitz [4] obtained a complete set of invariants under integral equivalence for "nice" local fields of characteristic $\neq 2$. The purpose of this note is to show the restriction on the characteristic is not necessary for either problem. The main results are contained in Theorems 1, 2, A, B, and C. We assume familiarity with [2].

1. Reduction of hermitian forms to quadratic forms.

Definition 1.1. E, $*$ is called a "nice" field if and only if one of the following conditions hold:

0). E is commutative, $\chi(E) \neq 2$, $*$ is the identity map of E, $F = E$.

1). E is commutative, $*$ is an automorphism of order 2, F is the fixed field of $*$. Hence, $[E:F] = 2$.

2). E is a reflexive division algebra with respective to $*$. F is the center of E. Hence, $[E:F] = 4$. If $\chi(E) \neq 2$, F is the set of fixed points of $*$. If $\chi(E) = 2$, the fixed points of $*$ has dimension 3 over F as a vector space. For details, see [3; p. 73, p. 84].

In all cases, for $A \in E$, $T(A) = A + A^* = T(A^*)$ and $N(A) = AA^* = N(A^*)$ both lie in F. $E^\#$ and $F^\#$ will denote the non-zero elements of E and F respectively.

Remark. In case 0), $\chi(E) = 2$ will lead to "symmetric matrices," which is not the same as quadratic forms. Similar results may be obtained. Since these are analogous to Section 4 and much simpler, we will exclude it from discussion.

Received February 21, 1963; revised received December 19, 1963.

* This work was prepared under NSF contract G-23833 at Harvard University.

Definition 1.2. Let V be a vector space over a "nice" field E, $*$. Then,

1). $\mathfrak{H}(V)$ is the set of all maps $\{\ ,\ \}: V \times V \to V$ such that $\{\ ,\ \}$ is E-linear in the first variable, $\{x, y\} = \{y, x\}^*$ and $\{x, x\} \in F$. The set $\mathfrak{H}(V)$ is an F-vector space by using the F-vector space structure of V. Thus $a \circ \{x, y\} = a\{x, y\}$ defines $a \circ \{\ ,\ \}$, where $a \in F$.

2). $\mathfrak{Q}^*(V)$ is the set of all maps $Q: V \to F$ such that

$$\langle x, y\rangle = Q(x + y) - Q(x) - Q(y)$$

is bilinear over F and $Q(Ax) = N(A)Q(x)$ for $A \in E$. As in 1) $\mathfrak{Q}^*(V)$ is an F-vector space. It is a subspace of the F-vector space of all quadratic maps on V, where V is considered only as an F-vector space. The conditions imply that $AA^*\langle Ax, y\rangle = \langle Ax, AA^*y\rangle = AA^*\langle x, A^*y\rangle$. Thus we have,

($*$). $\langle Ax, y\rangle = \langle x, A^*y\rangle$ for $x, y \in V$ and $A \in E$.

THEOREM 1. *Let E, $*$ be a nice field. V an E-vector space. For each $\{\ ,\ \} \in \mathfrak{H}(V)$, let $Q: V \to F$ be given by $Q(x) = \{x, x\}$. This then gives an F-isomorphism between $\mathfrak{H}(V)$ and $\mathfrak{Q}^*(V)$.*

Proof. Since E, $*$ is a nice field, the map $N: E \to F$ is a non-degenerate quadratic map on the F-vector space E. Let $u_i, v_j \in E$ be a pair of dual basis of E over F with respect to the associated inner product $\langle\ ,\ \rangle$. The construction of $\{x, y\}$ in terms of $\langle x, y\rangle$ is then given by $\{x, y\} = \sum_i \langle x, u_i y\rangle v_i$. This is the inverse of the map given in the theorem. Its verification follows from ($*$) and the fact that u_i, v_j are dual bases of E over F. Q. E. D.

(We are indebted to the referee for the above unified proof.)

Comparing Theorem 1 with the results of Arf [1], we have,

THEOREM 2. *Let E, $*$ be a "nice" field of characteristic 2 such that every quadratic space of dimension greater than 4 over F is isotropic. ($E.g.$, F is a global or a local field.) Then, every non-degenerate hermitian space V over E is E-equivalent to one of the following:*

1). $[E:F] = 4$, $V = \{1\} \oplus \cdots \oplus \{1\}$.
2). $[E:F] = 2$, $V = \{1\} \oplus \cdots \oplus \{1\} \oplus \{d\}$.

$\oplus$ is used to denote orthogonal sum, $\{a\}$ is an 1-dimensional hermitian space with basis x such that $\{x, x\} = a$. The quantity $d \neq 0$ determines a unique coset in $F^{\#}/N(E^{\#})$, called the discriminant of V.

2. Preparation to integral theory. From now on, we will assume that $E, *$ is a "nice" local field with $\chi(E)=2$. Thus $F=\Omega\langle\langle\pi\rangle\rangle$, where Ω is a finite field. $o=\Omega[[\pi]]$ and $O=$ integral closure of o in E. ord_E and ord_F are the ordinal functions in E and F respectively. $T(O)=\pi^s o$, s a non-negative integer. If E contains an unramified extension of degree 2 over F, then $s=0$. Conversely, if $s=0$, then by using Hensel's lemma, it follows that E contains an unramified extension of degree 2 over F. We will call $s=0$ the unramified case and $s>0$ the ramified case. U_E and U_F will denote the units of O and o respectively.

In any case $N(U_E)\subseteq U_F$. In the unramified case $N(U_E)=U_F$, [8; p. 90]. If $E, *$ is reflexive, then every commutative quadratic field extension of F may be embedded between F and E. This may be shown by using that every quadratic space of dimension >4 over F is isotropic and that under the quadratic map N, E is the canonical model of a non-isotropic quadratic space of dimension 4 over F. Thus this belongs to the unramified case and moreover $N(E^{\bullet})=F^{\bullet}$.

In the ramified case, we take $\Pi\in E$ such that $\mathrm{ord}_E(\Pi)=1$ and replace π by $N(\Pi)$.

LEMMA 2.1. *In the ramified case, $[E:F]=2$ and we have,*

1). $T(\Pi)=u\pi^s$, $\mathrm{ord}_F(u)=0$, $s>0$.

2). $T(\Pi^{2n}O)=T(\Pi^{2n+1}O)=\pi^{n+s}o$, *$n$ any integer.*

3). $1+\Pi^{2s}O\subseteq N(U_E)$.

4). *$U_F/N(U_E)$ is a cyclic group of order 2.*

5). *There is an additive map $\mathcal{P}'\colon\Omega\to\Omega$ with kernel of order 2 such that $N(1+\mu\Pi^{2s-1}+\cdots)=1+\mathcal{P}'(\mu)\pi^{2s-1}+\cdots$, where $\mu\in\Omega$. Thus we fix the choice of $\lambda\in\Omega$ such that $1+\lambda\pi^{2s-1}$ determines the non-trivial coset of $U_F/N(U_E)$.*

Proof. 1 and Π form an integral basis for O over o. Since T is linear over o and $\chi(E)=2$, we obtain 1) and 2). The rest follow from [8; p. 70, pp. 93-94].

If L is a free O-module of finite type furnished with a hermitian inner product, then L is called a hermitian lattice and the uniquely defined hermitian space $E\otimes_O L$ will be denoted by EL. We will always assume that L is non-degenerate. For $a\in F^{\bullet}$, $b,c\in F$ and $A\in E^{\bullet}$, $\{a\}$ and $\{{}_bA_c\}$ will denote the hermitian lattices with free bases x and y, z respectively such that $Q(x)=\{x,x\}=a$, and $Q(y)=b$, $\{y,z\}=A$, $Q(z)=c$. If the basis element is to be emphasized, we will write $Ox=\{a\}$, otherwise, $Ox\cong\{a\}$.

$s(L) = \{L, L\}$ is an O-module of finite type in E. $q(L) = \{{}_{\mathrm{o}}Q(x) \mid x \in L\}$ is an o-module of finite type in F. L is called i-modular if for each $x \in L$, $x \notin \Pi L$, we have $\{x, L\} = \Pi^i \mathrm{O}$. $L(i) = \{x \in L \mid \{x, L\} \subseteq \Pi^i \mathrm{O}\}$.

LEMMA 2.2. *Let L be a hermitian lattice and M an i-modular sublattice. Then M is an orthogonal summand of L if and only if $M \subseteq L(i)$. Thus L is an orthogonal sum of modular lattices each having dimension $\leqq 2$.*

Proof. Cf [6; Proposition 1, p. 160].

LEMMA 2.3. *Let L be a hermitian lattice. Then L admits a canonical decomposition of the form:*

$$L = \oplus L_i, \ \dim_{\mathrm{O}} L_i = r_i > 0, \ L_i \textit{ is } s(i)\textit{-modular}, \ s(1) < \cdots < s(m).$$

Any two such decompositions have the same type invariants r_i and $s(i)$.

Proof. Cf. [6; p. 161].

3. Unramified case.

THEOREM A. *Let E, * be unramified and L be a non-degenerate hermitian lattice. Then L admits an orthogonal decomposition of the form:*

$$L = \oplus L_i, \ \dim_{\mathrm{O}} L_i = r_i > 0, \ L_i = \{\pi^{s(i)}\} \oplus \cdots \oplus \{\pi^{s(i)}\}$$
$$s(1) < \cdots < s(m).$$

Thus the type invariants form a complete and independent set of invariants for L.

Proof. By Lemma 2.2 and Lemma 2.3 we may assume that L is i-modular of dimension $\leqq 2$.

If $\dim_{\mathrm{O}} L = 1$, then the assertion follows from $N(U_E) = U_F$. Thus let $\dim_{\mathrm{O}} L = 2$. From Lemma 2.2 and the case just considered, we may assume that $L = \mathrm{O}x + \mathrm{O}y = \{{}_a A_b\}$ with $\mathrm{ord}_E(A) < \mathrm{ord}_E(a) \leqq \mathrm{ord}_E(b)$. If $[E : F] = 4$, then modifying y by a unit, we may assume that $A = \mu \Pi^i$, $\Pi^2 = \pi$ and $\mu^2 + \mu = \mathfrak{P}(\mu) \in \Omega$ and $\mathfrak{P}(\mu) \notin \mathfrak{P}\Omega$. This follows from that E contains all possible quadratic extensions of F. Thus let $z = x + y$, we see that $\mathrm{O}z$ is an ortogonal summand of L. If $[E : F] = 2$, then modifying y by a unit, we may assume that $A = \mu\pi^i$, where $\mu^2 + \mu = \mathfrak{P}(\mu) \in \Omega$ and $\mathfrak{P}(\mu) \notin \mathfrak{P}\Omega$. Thus letting $z = x + y$ again shows that $\mathrm{O}z$ is an orthogonal summand of L. Q. E. D.

COROLLARY. *Let E, ‡ be unramified and L, M, N non-degenerate hermitian lattices. Then $L \oplus M \cong L \oplus N$ if and only if $M \cong N$.*

4. Ramified case.

Definition 4.1. We take Π as in Lemma 2.1. A hermitian lattice H is called an i-hyperbolic plane if it is isomorphic to $\{{}_0{}^{A}{}_0\}$ with $\mathrm{ord}_E(A) = i$. All i-hyperbolic planes are isomorphic. An orthogonal sum of i-hyperbolic planes is called an i-hyperbolic lattice.

LEMMA 4.2. *Let $L = H_j \oplus L_j$, $j = 1, 2$, be two decompositions of the hermitian lattice L. If H_j are both i-hyperbolic planes, then $L_1 \cong L_2$.*

Proof. From hypothesis given follows that $H_1 + H_2 \subseteq L(i)$ and $\{H_1, H_2\} \subseteq \Pi^i O$. If $\{H_1, H_2\} = \Pi^i O$, then we may proceed as in case 1 of the proof in [7; Theorem 3.3]. If $\{H_1, H_2\} \subseteq \Pi^{i+1} O$, then as in case 2 of the proof in [7; Theorem 3.3] we may assume that $L = H_1 + H_2$ and that L is i-modular. If $\dim_O L = 3$, then a calculation of the discriminant of L shows that $L_1 \cong L_2$. If $\dim_O L = 4$, then we proceed with the case 2 of the proof in [7; Theorem 3.3]. In all cases $\{\ ,\ \}$ and discriminant play the roles of $\langle\ ,\ \rangle$ and Arf invariant respectively.

LEMMA 4.3. *Let L be an i-modular plane. Then $Q(L) = q(L)$.*

Proof. By considering $a \circ L$, we may assume that $i = 0$ or 1 and $Q(x) = \pi^j$ where $q(L) = \pi^j O$. Thus, $2j \geqq i$.

If $2j = i$, then Lemma 2.2 implies that $L = Ox \oplus Oy = \{a\} \oplus \{b\}$. Thus by Lemma 2.1 and modularity, we may assume that $a = b = 1$. Since λ is a square in Ω and $\pi = N(\Pi)$, we see that $Q(x + By) = 1 + \lambda\pi^{2s-1}$ can be solved for $B \in O$. Thus from Lemma 2.1, we obtain $Q(L) = q(L)$.

If $2j > i$, then $j \geqq 1$. We can find $y \in L$ such that $\mathrm{ord}_E(\{x, y\}) = i$. Thus, $L = Ox + Oy = \{{}_a{}^{A}{}_b\}$, $\mathrm{ord}_E(A) = i < 2j = \mathrm{ord}_E(a) \leqq \mathrm{ord}_E(b)$, where $a = \pi^j$. By Lemma 2.1, we can find B with

$$\mathrm{ord}_E(B) \geqq 2j + 2s - 2 = 2(j + s - 1) > 0$$

such that $T(B^* A) = \pi^j(\lambda\pi^{2s-1})$. Thus

$$\mathrm{ord}_E(N(B)b) \geqq 2j + 2s - 2 + j \geqq 2s + j.$$

Hence,

$$Q(x + By) = \pi^j + T(B^* A) + N(B)b = \pi^j(1 + \lambda\pi^{2s-1} + c\pi^{2s}),\ \mathrm{ord}_F(c) \geqq 0.$$

By Lemma 2.1 again, we obtain $q(L) = Q(L)$. Q.E.D.

LEMMA 4.4. *Let L be an i-modular lattice with $\dim_O L \geqq 3$. Then, $L = M \oplus H$, where H is an i-hyperbolic plane.*

Proof. By Lemma 2.2, we may assume that $\dim_O L = 3$ or 4.

$\dim_{\mathrm{O}} L = 3$. By Theorem 1, the quadratic space associated to EL has dimension 6 over F, thus must be isotropic. Hence we may assume that $L = \mathrm{O}x \oplus (\mathrm{O}y + \mathrm{O}z) = \{a\} \oplus \{{}_b{}^A{}_0\}$, $\operatorname{ord}_E(a) = \operatorname{ord}_E(A) \leqq \operatorname{ord}_E(b)$. If $\operatorname{ord}_E(b) = i$, then by Lemma 2.2, $L \cong \{a\} \oplus \{b\} \oplus \{c\}$. By Lemma 2.1 we may assume that $a = b$. By Lemma 4.3 we may then assume that $a = b = c$. Taking the basis elements to be x, y and z, then $H = \mathrm{O}(x+y) + \mathrm{O}(y+z)$ is an i-hyperbolic plane. It is an orthogonal summand by Lemma 2.2. If $\operatorname{ord}_E(b) > i$, then by considering $a^{-1} \circ L$ and modifying y, z by units, we may assume that $a = A = 1$ and $b = \pi^j(1 + e\lambda\pi^{2s-1})$, $e = 0$ or 1, $0 < j \leqq \infty$. Replacing y by $y_1 = \pi^j x + y$, we obtain $\mathrm{O}y_1 + \mathrm{O}z = \{{}_c{}^1{}_0\}$, with $\operatorname{ord}_F(c) \geqq 2s$. By Lemma 2.1 we can find $C \in \mathrm{O}$ such that $Q(y_1 + Cz) = c + T(C) = 0$. Thus this sublattice is a desired hyperbolic plane.

$\dim_{\mathrm{O}} L = 4$. By the case just considered and Lemma 2.2, we may assume that $L = L_1 \oplus L_2$, $\dim_{\mathrm{O}} L_i = 2$ and $q(L) = \pi^j\mathfrak{o}$, $2j > i$. Again L is isotropic, we may therefore assume that $L_1 = \{{}_a{}^A{}_0\}$ and $L_2 = \{{}_b{}^A{}_c\}$. By Lemma 4.3 we may assume that $b\mathfrak{o} = q(L_2)$. If $\operatorname{ord}_F(a) \geqq \operatorname{ord}_F(b)$, then by Lemma 4.3, adding an element of L_2 to the first basis element of L_1 will give us an i-hyperbolic plane. Thus let $\operatorname{ord}_F(a) < \operatorname{ord}_F(b) \leqq \operatorname{ord}_F(c)$. Adding to the basis elements of L_2 suitable elements of ΠL_1 will give us a desired i-hyperbolic plane. Q. E. D.

LEMMA 4.5. *Let L be an i-modular plane, $i = 0$ or 1 and $q(L) = \pi^j\mathfrak{o}$. Then $j \leqq s$ and for $e = 0$ or 1, we have the following non-isomorphic normal forms for L:*

1). $i = 2j$, $L \cong \{1\} \oplus \{1 + e\lambda\pi^{2s-1}\}$.
2). $i < 2j$, $L \cong \{{}_a{}^A{}_b\}$, $a = \pi^j$, $A = \Pi^i$, $b = e\lambda\pi^{2s-1-j+i}$,

where $e = 0$ is understood in case $s = j$ and $i = 0$.

Furthermore, L is hyperbolic if and only if $j = s$ and $e = 0$.

Proof. By Lemma 4.3 we may assume that $L = \{{}_a{}^A{}_b\}$, where $a = \pi^j$, $A = \Pi^i$ and $\operatorname{ord}_F(b) \geqq j$. If $j > s$, then the discriminant of EL is $N(E^{*})$ by Lemma 2.1. Thus from Theorem 2, L is isotropic and we may assume that $b = 0$ and $\operatorname{ord}_F(a) \geqq j$. Letting x, y be the basis elements, we see that $Q(Bx + y) = N(B)\pi^j + T(B\Pi^i)$ will range over elements not in $\pi^{s+1}\mathfrak{o}$ by Lemma 2.1 and Lemma 4.3. Thus $j \leqq s$, a contradiction.

1) follows from Lemma 2.2 and Lemma 2.1.

Let $i < 2j$. If $s = j$ and $i = 0$, then we may proceed as in the proof of

$s \leqq j$ to obtain the desired assertion. Moreover, a similar argument will give us the last assertion. The details will be omitted.

Let $i < 2j$, $j \leqq k = \mathrm{ord}_F(b)$, where $L = \mathfrak{O}x + \mathfrak{O}y = \{{}_a{}^A{}_b\}$, x is fixed, $a = \pi^j$, $A = \Pi^i$ and y is a variable such that k is maximal. Suppose that $k < 2s - 1 - j + i$. Let $d = Q(Cx + y) = N(C)\pi^j + T(C\Pi^i) + b$. Since $\Omega^2 = \Omega$, we can find C with $\mathrm{ord}_E(C) = \mathrm{ord}_F(N(C)) = k - j \geqq 0$ such that $\mathrm{ord}_F(N(C)\pi^j + b) > k$. Since $i < 2j$, we obtain that $\mathrm{ord}_E(\{x, Cx + y\}) = \mathrm{ord}_E(C^*\pi^j + \Pi^i) = i$. Thus for suitable $B \in U_E$, we may replace b by $N(B)d$, this contradicts the maximality of k. Hence we may assume that $\mathrm{ord}_F(b) \geqq 2s - 1 - j + i$. By Lemma 2.1, we can find $B = 1 + C\Pi^{2s-1}$ with $\mathrm{ord}_E(C) \geqq 0$ such that $\pi^i(1 + \pi^{j-i}b) = \pi^i N(B)(1 + e\lambda\pi^{2s-1})$, $e = 0$ or 1. Consider the equation $\{x, Dx + By\} = D^*\pi^j + B^*\Pi^i = \Pi^i$. This may be solved with $\mathrm{ord}_E(D) = \mathrm{ord}_E(B + 1) + i - 2j \geqq 2s - 1 - 2j + i$. Since $s \geqq j$ and $i = 0, 1$, the exceptional case $s = j$, $i = 0$ has already been treated, we may assume that $\mathrm{ord}_E(D) \geqq 0$. Upon replacing y by $Dx + By$, the discriminant of L must be multiplied by the norm of the determinant of the transformation, i.e., by $N(B)$. Thus, by computing the discriminant of L, we see that b may be assumed to have the desired form. The discriminant computation together with Lemma 2.1 show that the different normal forms are not isomorphic. Q.E.D.

THEOREM B. *Let L and M be i-modular lattices, both non-degenerate. Then, $L \cong M$ if and only if $EL \cong EM$ and $q(L) = q(M)$.*

Proof. We only need to prove the sufficiency. From Theorem 2, Lemma 4.2, 4.4, we may assume that $\dim_{\mathfrak{O}} L = \dim_{\mathfrak{O}} M = 2$. After scaling, i.e., by considering $a \circ L$ and $a \circ M$, we may assume that $i = 0$, of 1. The assertion now follows from Lemma 4.5. Q.E.D.

Definition 4.6. Let L be non-degenerate hermitian lattice, then $D(L)$ will denote the coset in $F^*/N(U_E)$ determined by the discriminant of L with respect to an integral basis of L. If $S \subseteq F$, then

$$\mathrm{Ord}_F(S) = \mathrm{l.u.b.}\{\mathrm{ord}_F(s) \mid s \in S\}.$$

THEOREM C. *Let $L = \oplus L_i$ and $M = \oplus M_j$, $1 \leqq i \leqq m$, $1 \leqq j \leqq n$, be canonical decompositions of the non-degenerate hermitian lattices L and M respectively. Then $L \cong M$ if and only if the following hold:*

1). *$EL \cong EM$.*

2). *L and M have the same type, i.e., $m = n$, L_i and M_i are $s(i)$-modular, $\dim_{\mathfrak{O}} L_i = \dim_{\mathfrak{O}} M_i = r_i$ and $s(1) < \cdots < s(m)$.*

3). *$q(L(i) \oplus H_i) = q(M(i) \oplus H_i) = \pi^{u(i)}\mathfrak{o}$, H_i an i-hyperbolic plane, $s(1) \leqq i \leqq s(m)$. Thus, $2u(i) \geqq i$ and $u(i+1) \geqq u(i)$.*

4). *$\mathrm{Ord}_F(D_i(L)/D_i(M) + 1) \geqq u(i) + u(i+1) - i (> 0)$, $s(1) \leqq i < s(m)$, where $D_i(L) = D(L_1 \oplus \cdots \oplus L_j)$ with $s(j) \leqq i < s(j+1)$, similarly for $D_i(M)$.*

Proof. By Lemma 4.2, we see that our conditions are "invariant" under the addition of hyperbolic planes simultaneously to L and M. Thus, we may assume that $s(i+1) = s(i) + 1$ and $r_i \geqq 3$. By Lemma 2.2, Lemma 4.5 and Theorem B, we may assume that L_i and M_i have been broken down to orthogonal sums of lines or planes according to r_i is odd or even. If L is replaced by $a \circ L$, $a \in F^{\#}$, then $s(i)$ changes to $s(i) + 2 \cdot \mathrm{ord}_F(a)$ and $u(i)$ changes to $u(i) + \mathrm{ord}_F(a)$. Thus the conditions are also "invariant" under this scaling operation.

Necessity. We only need to check 4). Using an argument similar to [7; Theorem 2.3, Lemma 5.4], we may assume that the two decompositions differ only in adjacent components each of dimension at most 2. Thus we may assume that $m = 2$, $r_i \leqq 2$. If $L_1 = \{a\}$, then we only need to consider the change $\{a\} \to \{a + b\}$, where $b \in q(L_2)$. Thus the discriminant changes by a factor $1 + a^{-1}b$. But $\mathrm{ord}_F(a^{-1}b) \geqq u(s(2)) - u(s(1))$, $s(1) = 2u(s(1))$. Thus we obtain 4), since the right side is positive implies that 4) is an equivalence relation. If $L_1 = \{{}_{a}{}^{A}{}_{b}\}$, then we only need to consider the change $\{{}_{a}{}^{A}{}_{b}\} \to \{{}_{a}{}^{A}{}_{b+c}\}$, where $c \in q(L_2)$. If $2u(s(1)) = s(1)$, then by Lemma 4.4, we can break up L_1 and apply the preceding case. Thus let $2u(s(1)) > s(1)$. The ratio of the discriminants is

$$1 + ac(N(A) + ab)^{-1}, \ \mathrm{ord}_F(ac(N(A) + ab)^{-1}) \geqq u(s(1)) + u(s(2)) - s(1).$$

Since the right side is positive, 4) is an equivalence relation, hence we obtain 4).

Sufficiency. The case $m = 1$ follows from Theorem B. An argument similar to [7; Lemma 5.3] may be carried out, using Lemma 4.3 and Lemma 4.4, to allow us to assume that L and M are saturated in each component, i.e., $q(L_i) = q(M_i) = \pi^{u(s(i))}\mathfrak{o}$, $1 \leqq i \leqq m$. By Theorem B and the necessity, we may assume that $L_i = K_i \oplus H_i$, where $\dim_{\mathfrak{O}} H_i = 2$, $q(K_i) = q(H_i) = q(L_i)$, $D(EK_1) = 1$, $D(EH_2) = 1$. Thus we may assume that $H_1 = \{{}_{a}{}^{A}{}_{b}\}$, $a = \pi^{u(s(1))}$, $A = \Pi^{s(1)}$ and $H_2 = \{{}_{c}{}^{B}{}_{0}\}$, $\mathrm{ord}_F(c) = u(s(2))$, $\mathrm{ord}_B(B) = s(2) = s(1) + 1$. By adding a suitable element of H_2 to the second basis element of H_1 and using Lemma 4.3, we may alter b by an arbitrary element of $\pi^{u(s(2))}\mathfrak{o}$. This transformation in $H_1 \oplus H_2$ will not destroy the saturation, but will, by 4), allow us to assume that $D(L_1) = D(M_1)$. Thus Theorem B together with 2) and 3) will allow us to assume that $L_1 \cong M_1$. The four conditions given will then imply a similar set of four conditions after we cancel out L_1 and M_1. Hence by induction, we obtain that $L \cong M$. Q. E. D.

HARVARD UNIVERSITY,
UNIVERSITY OF PENNSYLVANIA.

REFERENCES.

[1] C. Arf, "Untersuchungen über quadratische Formen in Körpern der Charakteristik 2," *Journal für die reine und angewandte Mathematik*, vol. 183 (1940), pp. 148-167.
[2] N. Bourbaki, "Algèbre," Chapitre 9, *Actualités Scientifiques et Industrielles*, No. 1272, Paris, 1959.
[3] J. Dieudonné, *Sur les groupes classiques*, Paris, 1948.
[4] R. Jacobwitz, "Hermitian forms over local fields," *American Journal of Mathematics*, vol. 84 (1962), pp. 441-465.
[5] N. Jacobson, "A note on Hermitian forms," *Bulletin of the American Mathematical Society*, vol. 46 (1940), pp. 264-268.
[6] O. T. O'Meara, "Integral equivalence of quadratic forms in ramified local fields," *American Journal of Mathematics*, vol. 79 (1957), pp. 157-186.
[7] C. H. Sah, "Quadratic forms over fields of characteristic 2," *American Journal of Mathematics*, vol. 82 (1960), pp. 812-830.
[8] J.-P. Serre, *Corps locaux*, Paris, 1962.

INVARIANT DISTRIBUTIONS ON LIE ALGEBRAS.*

By HARISH-CHANDRA.

1. Introduction. The object of this paper is to start laying the groundwork for the proof of Theorem 1 of [4(e)], which says, roughly, that the irreducible characters of a semisimple Lie group, which had so far only been known to be distributions [4(b)], are actually functions. Our method consists in splitting the problem into two parts. The first one deals with an analogous question on the Lie algebra (see Lemma 1 of [4(e)]) and the second with the task of lifting this result from the algebra to the group itself. In this paper we shall be mainly concerned with one of the principal steps [4(e), Lemma 2] needed in the first part.

Our entire approach to the problem is based on a simple result (Theorem 1) about a C^∞ mapping π of a manifold M *onto* another manifold N, which has the maximum rank everywhere.[1] This allows one to transfer a distribution from N to M. Consider the following example. Let G be the orthogonal group of degree n operating in the usual way on $\boldsymbol{R}^n$. Identify $\boldsymbol{R}$ with a subspace of $\boldsymbol{R}^n$ under the mapping $t \rightarrow (t, 0, 0, \cdots, 0)$ $(t \in \boldsymbol{R})$. Let $\boldsymbol{R}_+$ be the set of all positive real numbers. Now let $M = G \times \boldsymbol{R}_+$, N the complement of the origin in $\boldsymbol{R}^n$ and π the mapping $(x, r) \rightarrow xr$ $(x \in G,\ r \in \boldsymbol{R}_+)$ of M onto N. Let f be a C^∞ function on N which is invariant under G. Then $f(xr)$ is independent of x. So in this way we get a function σ_f on $\boldsymbol{R}_+$. D being a differential operator on N which is invariant under G, one can show that there exists a differential operator $\Delta(D)$ on $\boldsymbol{R}_+$ such that $\sigma_{Df} = \Delta(D)\sigma_f$ for any invariant C^∞ function f on N. Theorem 1 enables us to prove similar results about distributions (see Theorems 2 and 3).

The most significant result of this paper is contained in Theorem 4. This will be used in subsequent papers to prove the theorem about characters, mentioned above.

The contents of this paper are as follows. In § 2 we collect some well-known facts about manifolds, differential operators and distributions. § 3 is devoted to the proof of Theorem 1, which, as we have just indicated, is rather basic for our purposes. In § 5 we introduce invariant differential operators and distributions and consider some of their elementary properties. The

---

* Received October 25, 1963.

[1] I understand from Prof. L. Schwartz that he had independently obtained a similar result some years ago.

results of § 7 are important for applications. They enable us to construct the above-mentioned mapping $D \to \Delta(D)$. The main result concerning $\Delta(D)$ and the invariant functions, is proved in Lemma 10. In Theorems 2 and 3 (of § 9), this is extended to invariant distributions. This general theory is then applied to the case of a semisimple Lie algebra to prove Theorem 4. Our argument depends on a detailed consideration of some differential operators associated with the Killing form (see Lemmas 30 and 31).

Lemma 16 was pointed out to me by Borel who also gave me the proof of Lemma 17. These lemmas play an essential role in the proof of Theorem 4.

This work was supported, in part, by a grant from the Sloan Foundation and a contract with the U. S. Army.

2. Differential operators and distributions. By a manifold M we mean a Hausdorff and locally Euclidean space satisfying the second axiom of countability, all of whose connected components have the same dimension m and are endowed with a C^∞ structure. It is oriented if each of its connected components is oriented. Let $C^\infty(M)$ denote the space of all C^∞ functions from M to $\mathbf{C}$ and $\mathcal{E}$ the space of all $\mathbf{C}$-linear mappings of $C^\infty(M)$ into itself. Then $\mathcal{E}$ is an associative algebra. If X is a (C^∞) vector-field on M and $g \in C^\infty(M)$, the mappings $E_X\colon f \to Xf$ and $E_g\colon f \to gf$ $(f \in C^\infty(M))$ are both in $\mathcal{E}$. Let $\mathfrak{D} = \mathfrak{D}(M)$ denote the subalgebra of $\mathcal{E}$ generated by E_X and E_g for all X and g. Then $\mathfrak{D}$ is called the algebra of differential operators on M. It would be convenient to denote E_X and E_g by X and g respectively. However, if it is necessary to avoid confusion, we shall write the product of two differential operators D_1, D_2 as $D_1 \circ D_2$. Thus $E_X E_g = X \circ g$ while $E_{Xg} = Xg$. If $D \in \mathfrak{D}$ and $f \in C^\infty(M)$, we shall often denote the value of Df at a point $p \in M$ by $f(p;D)$.

For any $D \in \mathfrak{D}$, $\operatorname{Supp} D$ is the closed subset of M defined as follows. Let $p \in M$. Then $p \notin \operatorname{Supp} D$ if and only if $p \notin \operatorname{Supp} Df$ for every $f \in C^\infty(M)$. An indexed family $\{F_\alpha\}_{\alpha \in A}$ of sets in M is called scattered if, for every compact subset K of M, $K \cap F_\alpha = \emptyset$ for all except a finite number of α in A. A subset Φ of $\mathfrak{D}$ is scattered if the family $\{\operatorname{Supp} D\}_{D \in \Phi}$ is scattered. Let $C_c^\infty(M)$ denote the subspace of all functions in $C^\infty(M)$ whose support is compact. Put $\nu_D(f) = \sup_M |Df|$ $(D \in \mathfrak{D})$ and for any scattered subset Φ of $\mathfrak{D}$, define $\nu_\Phi(f) = \sum_{D \in \Phi} \nu_D(f)$ $(f \in C_c^\infty(M))$. Let S be the class of all scattered subsets of $\mathfrak{D}$. Then the seminorms $\{\nu_\Phi; \Phi \in S\}$ define the structure of a locally convex space on $C_c^\infty(M)$. The corresponding topology is called the Schwartz topology. K being any compact subset of M, let $C_K^\infty(M)$ denote the subspace of all functions in $C_c^\infty(M)$ whose support is contained in K. It is easy to see that the collections $\{\nu_D; D \in \mathfrak{D}\}$ and $\{\nu_\Phi; \Phi \in S\}$ define the same topology on $C_K^\infty(M)$. U being an open subset of M, any function in $C_c^\infty(U)$ can be extended to a function on M by setting it zero outside U. In this way $C_c^\infty(U)$ becomes a subspace of $C_c^\infty(M)$ and it is easy to see that the injection $C_c^\infty(U) \to C_c^\infty(M)$ is continuous.

A distribution T on M is a continuous linear mapping of $C_c^\infty(M)$ into $\boldsymbol{C}$.

LEMMA 1. *Let V be a complex locally convex space and T a linear mapping of $C_c^\infty(M)$ into V. Assume that T is continuous on $C_K^\infty(M)$ for every compact subset K of M. Then T is continuous.*

This is an immediate consequence of Theorem II (p. 69) of [9].

LEMMA 2. *Let T be a linear mapping of $C_c^\infty(M)$ into $\boldsymbol{C}$. Then it is a distribution if and only if it satisfies the following condition. Given any open and relatively compact subset U of M, we can choose a finite number of differential operators $D_1, \cdots, D_r$ on M such that*

$$|T(f)| \leqq \sum_{1 \leqq i \leqq r} \sup |D_i f| \qquad (f \in C_c^\infty(U)).$$

This is well known and is an easy consequence of Lemma 1.

Let M_1, M_2 be two manifolds and $M_1 \times M_2$ their Cartesian product. For $f_i \in C_c^\infty(M_i)$ $(i = 1, 2)$ let $f_1 \times f_2$ denote the function $(p_1, p_2) \to f_1(p_1) f_2(p_2)$ $(p_i \in M_i;\ i = 1, 2)$ on $M_1 \times M_2$. In this way we get a linear mapping of $C_c^\infty(M_1) \otimes C_c^\infty(M_2)$ into $C_c^\infty(M_1 \times M_2)$.

LEMMA 3. *The image of $C_c^\infty(M_1) \otimes C_c^\infty(M_2)$ under this mapping is dense in $C_c^\infty(M_1 \times M_2)$.*

This is well known (see [9, p. 107, Theorem III]).

It follows from Lemma 3 that given $D_i \in \mathfrak{D}(M_i)$ $(i = 1, 2)$, there exists exactly one differential operator D on $M_1 \times M_2$ such that $D(f_1 \times f_2) = D_1 f_1 \times D_2 f_2$ $(f_i \in C_c^\infty(M_i), i = 1, 2)$. We denote D by $D_1 \times D_2$.

3. **Proof of Theorem 1.** Let ω be a (C^∞) differential form on M of degree m. Fix a point $p \in M$ and let $(x_1, \cdots, x_m)$ be a coordinate system on some open and connected neighborhood U of p in M. Then there exists a unique function $\alpha \in C^\infty(U)$ such that $\omega = \alpha\, dx_1 \wedge dx_2 \wedge \cdots \wedge dx_m$ on U. We say $\omega \neq 0$ at p if $\alpha(p) \neq 0$. Moreover if M is oriented and ω is real, we say $\omega > 0$ at p if $\epsilon\alpha(p) > 0$. Here $\epsilon = 1$ or -1 according as the coordinate system $(x_1, \cdots, x_m)$ is positive or negative with respect to the given orientation of M. We write $\omega > 0$ if ω is positive everywhere.

Now suppose $\omega > 0$. Then it defines a (positive) Borel measure μ on[2] M. For example if f is a continuous function on M with compact support, we have

$$\int f d\mu = \int_M f\omega.$$

Let M and N be two oriented manifolds of dimensions m and n respectively and $\pi: M \rightarrow N$ a surjective C^∞ mapping such that rank $d\pi = n$ everywhere.

THEOREM 1.[1] *Let ω_M and ω_N be two differential forms on M and N respectively of degree m and n. Assume that ω_N is nowhere zero. Then for every $\alpha \in C_c^\infty(M)$, there exists a unique function $f_\alpha \in C_c^\infty(N)$ such that*

$$\int_M (F \circ \pi) \alpha\omega_M = \int_N F f_\alpha \omega_N$$

for all $F \in C_c^\infty(N)$. Moreover $\operatorname{Supp} f_\alpha \subset \pi(\operatorname{Supp} \alpha)$ *and $\alpha \rightarrow f_\alpha$ is a continuous mapping of $C_c^\infty(M)$ into $C_c^\infty(N)$. Finally if ω_M is nowhere zero, this mapping is surjective.*

COROLLARY 1. *Let F be a continuous function on N. Then*

$$\int_M (F \circ \pi) \alpha\omega_M = \int_N F f_\alpha \omega_N \qquad (\alpha \in C_c^\infty(M)).$$

COROLLARY 2. *Assume $\omega_M > 0$, $\omega_N > 0$ and let μ_M and μ_N denote the corresponding Borel measures on M and N respectively. Suppose F is a Borel measurable function on N. Then F is locally summable (with respect to μ_N) if and only if $F \circ \pi$ is locally summable (with respect to μ_M) and, in case this is so, we have*

$$\int_M (F \circ \pi) \alpha \, d\mu_M = \int_N F f_\alpha \, d\mu_N \qquad (\alpha \in C_c^\infty(M)).$$

Since ω_N is nowhere zero, the uniqueness of f_α in Theorem 1 follows from its definition. Hence we are mainly concerned with its existence. We shall first prove Theorem 1 and its corollaries in a special case.

For any $a > 0$, let I_a denote the open interval $(-a, a)$ in $\boldsymbol{R}$ and J_a the closed interval $[-a, a]$. Let π denote the projection $(x_1, \cdots, x_m) \rightarrow (x_1, \cdots, x_n)$ of $\boldsymbol{R}^m$ on $\boldsymbol{R}^n$.

---

[2] It is convenient to express this relationship by writing $\omega \sim \mu$. Similarly in other cases.

Lemma 4. *For any* $\alpha \in C_c^\infty(I_a^m)$ *define* $g_\alpha \in C_c^\infty(I_a^n)$ *by*

$$g_\alpha(x_1, \cdots, x_n) = \int \alpha(x_1, \cdots, x_m)\, dx_{m+1} \cdots dx_n.$$

Then $\alpha \rightarrow g_\alpha$ *is a continuous mapping of* $C_c^\infty(I_a^m)$ *onto* $C_c^\infty(I_a^n)$ *and* $\operatorname{Supp} g_\alpha \subset \pi(\operatorname{Supp} \alpha)$.

It is obvious that $g_\alpha \in C_c^\infty(I_a^n)$ and $\operatorname{Supp} g_\alpha \subset \pi(\operatorname{Supp} \alpha)$. For any b $(0 < b < a)$, put $C_b^\infty(I_a^m) = C_K^\infty(I_a^m)$ where $K = J_b^m$. Then it is easy to verify that the mapping is continuous on $C_b^\infty(I_a^m)$. Therefore we conclude from Lemma 1 that it is also continuous on $C_c^\infty(I_a^m)$. So it only remains to show that it is surjective. For this we may assume that $m > n$. Fix $\gamma \in C_c^\infty(I_a^{m-n})$ such that

$$\int \gamma(x_{n+1}, \cdots, x_m)\, dx_{n+1} \cdots dx_m = 1.$$

Then for a given $\beta \in C_c^\infty(I_a^n)$, put $\alpha = \beta \times \gamma$. Then $\alpha \in C_c^\infty(I_a^m)$ and $\beta = g_\alpha$.

Now in Theorem 1 first consider the case when $M = I_a^m$, $N = I_a^n$ and π is the above projection. Let

$$dx = dx_1 \wedge dx_2 \wedge \cdots \wedge dx_m \text{ and } dy = dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n.$$

Then $\omega_M = \gamma_M\, dx$ and $\omega_N = \gamma_N\, dy$ where γ_M and γ_N are C^∞ functions on M and N respectively and γ_N is nowhere zero. For any $\alpha \in C_c^\infty(M)$, define f_α on N as follows:

$$f_\alpha = \epsilon_M \epsilon_N \gamma_N^{-1} g_{\gamma_M \alpha}.$$

Here ϵ_M, ϵ_N are constants equal to ± 1 chosen in such a way that $\epsilon_M\, dx > 0$ on M and $\epsilon_N\, dy > 0$ on N. Then f_α obviously fulfills the condition of Theorem 1. Moreover it follows from Lemma 4 that $\alpha \rightarrow f_\alpha$ is a continuous mapping of $C_c^\infty(M)$ into $C_c^\infty(N)$ and $\operatorname{Supp} f_\alpha \subset \pi(\operatorname{Supp} \alpha)$. Now assume ω_M is nowhere zero. Then the endomorphisms $\alpha \rightarrow \gamma_M \alpha$ and $g \rightarrow \gamma_N^{-1} g$ of $C_c^\infty(M)$ and $C_c^\infty(N)$ respectively are surjective. Hence we conclude from Lemma 4 that the mapping $\alpha \rightarrow f_\alpha$ is surjective. Finally Corollaries 1 and 2 are, in this case, immediate consequences of Fubini's Theorem.

We come now to the general case. Fix a point $p \in M$ and choose a small open connected neighborhood U_0 of p in M. Let $V_0 = \pi(U_0)$. Since rank $d\pi = n$ everywhere, V_0 is open in N and if U_0 is sufficiently small, we can choose a coordinate system $(x_1, \cdots, x_m)$ on U_0 such that $x_i(p) = 0$ $(1 \leqq i \leqq m)$ and $x_j = y_j \circ \pi$ $(1 \leqq j \leqq n)$ where $(y_1, \cdots, y_n)$ is a coordinate system on V_0. Fix a small number $a > 0$ and let U be the set of all $q \in U_0$ where $|x_i(q)| < a$ $(1 \leqq i \leqq m)$. If a is sufficiently small, $q \rightarrow (x_1(q), \cdots, x_m(q))$

is a diffeomorphism of U with $I_a{}^m$. Put $V=\pi(U)$. Then it is clear from our discussion above that Theorem 1 and its corollaries hold for (U,V,π). This shows that for every $p\in M$, we can choose an open neighborhood U_p of p in M such that Theorem 1 and its corollaries hold for $(U_p,\pi(U_p),\pi)$.

Fix a compact subset K of M. Then $\{U_p\}_{p\in K}$ is an open covering of K in M. We can therefore select a finite subcovering $\{U_i\}_{1\leqq i\leqq r}$ from it. Choose $\beta_i\in C_c{}^\infty(U_i)$ such that $\beta_1+\cdots+\beta_r=1$ on K. Put $V_i=\pi(U_i)$ and for any $\alpha\in C_K{}^\infty(M)$, define $\alpha_i=\beta_i\alpha$. Then $\alpha=\alpha_1+\cdots+\alpha_r$ and $\alpha_i\in C_c{}^\infty(U_i)$. Let $f_{\alpha_i}\in C_c{}^\infty(V_i)$ be the function which corresponds to α_i under Theorem 1 for (U_i,V_i,π). Put $f_\alpha=\sum_{1\leqq i\leqq r}f_{\alpha_i}$. Then if F is a continuous function on N,

$$\int (F\circ\pi)\alpha\omega_M=\sum_{1\leqq i\leqq r}\int_M (F\circ\pi)\alpha_i\omega_M.$$

But

$$\int_M (F\circ\pi)\alpha_i\omega_M=\int_{U_i}(F\circ\pi)\alpha_i\omega_M=\int_{V_i}Ff_{\alpha_i}\omega_N=\int_N Ff_{\alpha_i}\omega_N$$

from Corollary 1 applied to (U_i,V_i,π). Hence

$$\int_M (F\circ\pi)\alpha\omega_M=\sum_i\int_N Ff_{\alpha_i}\omega_N=\int Ff_\alpha\omega_N.$$

Moreover $\operatorname{Supp}f_\alpha\subset\bigcup_i\operatorname{Supp}f_{\alpha_i}$. But $\operatorname{Supp}f_{\alpha_i}\subset\pi(\operatorname{Supp}\alpha_i)\subset\pi(\operatorname{Supp}\alpha)$ since Theorem 1 holds for (U_i,V_i,π). Hence $\operatorname{Supp}f_\alpha\subset\pi(\operatorname{Supp}\alpha)$ and the mapping $\alpha\to f_\alpha$ is continuous on $C_K{}^\infty(M)$. Now assume ω_M is nowhere zero and fix $g\in C_c{}^\infty(N)$. Then we can select the compact set K above in such a way that $\operatorname{Supp}g\subset\pi(K)$. Now $\{V_i\}_{1\leqq i\leqq r}$ is an open covering for $\pi(K)$ in N. Choose $\rho_i\in C_c{}^\infty(V_i)$ such that $\rho_1+\rho_2+\cdots+\rho_r=1$ on $\pi(K)$. Put $g_i=\rho_i g$. Then $g=g_1+\cdots+g_r$. Since Theorem 1 is applicable to (U_i,V_i,π), we can choose $\alpha_i\in C_c{}^\infty(U_i)$ such that $g_i=f_{\alpha_i}$. Then if $\alpha=\alpha_1+\cdots+\alpha_r$, it is obvious that $g=f_\alpha$. This proves Theorem 1 and Corollary 1 above.

Now we come to Corollary 2. Fix a point $q\in N$ and a point $p\in M$ such that $\pi(p)=q$. Select U_p as above and define $V=\pi(U_p)$. Then Corollary 2 holds for (U_p,V,π) and therefore $F\circ\pi$ is locally summable on V if and only if F is locally summable on U_p. Since q was any point of N, the first statement of Corollary 2 follows. Now assume that F is locally summable. Then if $\alpha\in C_c{}^\infty(M)$,

$$\begin{aligned}\int_M (F\circ\pi)\alpha d\mu_M&=\sum_i\int_{U_i}(F\circ\pi)\alpha_i d\mu_M\\&=\sum_i\int_{V_i}Ff_{\alpha_i}d\mu_N=\int_N Ff_\alpha d\mu_N\end{aligned}$$

in the above notation.

We keep to the notation of Theorem 1.

LEMMA 5. *Let T be a distribution on N. Then the mapping $\tau: \alpha \rightarrow T(f_\alpha)$ $(\alpha \in C_c^\infty(M))$ is a distribution on M. Moreover if ω_M is nowhere zero, T is completely determined by τ.*

The first assertion follows from the continuity of the mapping $\alpha \rightarrow f_\alpha$ and the second from its surjectivity (when ω_M is nowhere zero).

4. Adjoint of a differential operator. Assume that M is oriented and fix, once for all, a differential form $\omega > 0$. Then for every $D \in \mathfrak{D}(M)$, there exists a unique differential operator D^*, called the adjoint of D (with respect to ω), such that

$$\int_M Df \cdot \phi\omega = \int_M f \cdot D^*\phi \cdot \omega$$

whenever f and ϕ are two C^∞ functions on M, at least one of which has compact support. The mapping $D \rightarrow D^*$ is an anti-automorphism of $\mathfrak{D}(M)$ of order 2.

Let $\mathfrak{T}$ be the space of all distributions on M. Then we turn $\mathfrak{T}$ into a $\mathfrak{D}(M)$-module as follows. If $T \in \mathfrak{T}$ and $D \in \mathfrak{D}(M)$, the mapping $f \rightarrow T(D^*f)$ $(f \in C_c^\infty(M))$ is a distribution which, by definition, is DT. Let μ be the Borel measure on M corresponding to ω (see § 3). If F is a function on M which is locally summable (with respect to μ), the mapping

$$T_F: f \rightarrow \int fF\, d\mu \qquad (f \in C_c^\infty(M))$$

is a distribution. If we identify two locally summable functions whose difference is almost everywhere zero, then the mapping $F \rightarrow T_F$ is injective. Let T be a distribution on M. Then we say that T is a function if $T = T_F$ for some locally summable function F. In that case we also write $T = F$.

If each connected component of M is endowed with an analytic structure, we say that M is an analytic manifold. Suppose M is analytic and D is a differential operator on M. Then D is analytic if for any analytic function f on an open subset U of M, the function Df is also analytic on U. Assume that the differential form ω above is analytic. Then if D is analytic, so is also D^*.

5. Invariant differential operators and distributions. Let G be a group and M a manifold. Assume that G operates on M. This means that for every $g \in G$, there is given a diffeomorphism $x \rightarrow x^g$ $(x \in M)$ of M such that $x^1 = x$ and $(x^{g_2})^{g_1} = x^{g_1 g_2}$ $(g_1, g_2 \in G;\ x \in M)$. Let Ω_r denote the space of all (C^∞) differential forms on M of degree r. Then G operates linearly on $C^\infty(M)$, $\mathfrak{D}(M)$ and Ω_r as follows. Fix $g \in G$. Then

$$f^g(x) = f(x^{g^{-1}}) \qquad (f \in C^\infty(M), x \in M),$$
$$D^g f = (Df^{g^{-1}})^g \qquad (D \in \mathfrak{D}(M), f \in C^\infty(M)),$$
$$\omega^g(X_1, \cdots, X_r) = (\omega(X_1^{g^{-1}}, \cdots, X_r^{g^{-1}}))^g \qquad (\omega \in \Omega_r)$$

where $X_1, \cdots, X_r$ are any r vector-fields on M. It is obvious that $f \to f^g$ ($f \in C_c^\infty(M)$) is a continuous endomorphism of $C_c^\infty(M)$. Therefore if T is a distribution on M, the mapping $T^g\colon f \to T(f^{g^{-1}})$ is also a distribution. In particular if μ is a measure on M, the same holds for μ^g.

Now assume M is oriented and we have fixed ω as in § 4. Let μ be the measure on M corresponding to ω (see § 3). Assume that $\mu^g = \mu$ for all $g \in G$ and fix $D \in \mathfrak{D}(M)$. Then it is easy to verify that $(D^g)^* = (D^*)^g$ and $(DT)^g = D^g T^g$ ($g \in G$) for any distribution T on M. Let F be a locally summable function on M. Then if $T = F$, it follows that $T^g = F^g$ where F^g is the function $x \to F(x^{g^{-1}})$ ($x \in M$).

A distribution T is called invariant, if $T^g = T$ for all $g \in G$. Similarly a differential operator D is called invariant if $D^g = D$ ($g \in G$).

Actually, we shall mostly be concerned with the case when G is a Lie group and the mapping $(g, x) \to x^g$ of $G \times M$ into M is C^∞. Let $\mathfrak{g}$ be the Lie algebra of G. Then for any $X \in \mathfrak{g}$, we define a vector-field $\tau(X)$ on M as follows:

$$f(x; \tau(X)) = \{df^{\exp tX}(x)/dt\}_{t=0} \qquad (f \in C^\infty(M), x \in M).$$

It is easy to verify that $[\tau(X), \tau(Y)] = \tau([X, Y])$ ($X, Y \in \mathfrak{g}$). A distribution T on an open subset U of M is called locally invariant if $\tau(X)T = 0$ for all $X \in \mathfrak{g}$. Similarly a differential operator D on U is said to be locally invariant if $\tau(X) \circ D = D \circ \tau(X)$ for $X \in \mathfrak{g}$.

LEMMA 6. $\tau(X)^* = -\tau(X)$ ($X \in \mathfrak{g}$) *and every invariant distribution (or differential operator) on M is locally invariant.*

Fix $f, \phi \in C_c^\infty(M)$. Then

$$\int_M f^g \cdot \phi \cdot \omega = \int_M f \cdot \phi^{g^{-1}} \cdot \omega \qquad (g \in G).$$

Hence if $X \in \mathfrak{g}$,

$$\int_M \tau(X)f \cdot \phi\omega = \{(d/dt) \int_M f^{\exp tX} \cdot \phi \cdot \omega\}_{t=0}$$
$$= \{(d/dt) \int_M f\phi^{\exp(-tX)} \cdot \omega\}_{t=0} = -\int_M f \cdot \tau(X)\phi \cdot \omega.$$

This proves that $\tau(X)^* = -\tau(X)$. Moreover it is easy to verify that

$$t^{-1}(f^{\exp tX} - f) \to \tau(X)f$$

in $C_c^\infty(M)$ as $t \to 0$. From this the assertion about an invariant distribution follows immediately. Let D be an invariant differential operator. Then if $\exp tX = x(t)$, we have

$$(Df)^{x(t)} = Df^{x(t)}.$$

Since D defines a continuous endomorphism of $C_c^\infty(M)$, it is clear that

$$\begin{aligned}\tau(X)(Df) &= \lim_{t\to 0} t^{-1}\{(Df)^{x(t)} - Df\}\\ &= D(\tau(X)f).\end{aligned}$$

This proves that D is locally invariant.

We now consider an example. Take $M = G$ and define $x^g = gx$ $(g, x \in G)$ and let dx denote the left-invariant Haar measure on G and $\omega > 0$ the corresponding left-invariant differential form. We write $X' = \tau(X)$ $(X \in \mathfrak{g})$ in this case so that

$$f(x; X') = -\{df(\exp tX \cdot x)/dt\}_{t=0} \qquad (f \in C^\infty(G), x \in G).$$

Lemma 7. *Let T be a distribution on an open connected subset U of G. Suppose $X'T = 0$ for all $X \in \mathfrak{g}$. Then T is a constant.*

Select a base $X_1, \cdots, X_n$ for $\mathfrak{g}$ over $\boldsymbol{R}$ and put $D = X'^2_1 + \cdots + X'^2_n$. Then obviously D is an elliptic differential operator on G. Therefore since $DT = 0$, we can conclude that there exists a C^∞ function F on U such that $T = F$. But then $X'F = 0$ for all X in $\mathfrak{g}$. Since U is connected, this implies that F is a constant.

6. Some notations and elementary facts. Let E be a vector space of finite dimension over an infinite field k and $P(E)$ the algebra of all polynomial functions from E to k. Let E' be the space of all linear functions on E. Define $S(E) = P(E')$. We call $S(E)$ and $P(E)$ respectively the symmetric and polynomial algebras over E. Since E is the dual of E', $E \subset S(E)$. Let F be another finite-dimensional vector space over k and $i\colon F \to E$ a linear mapping. Then i can be extended uniquely to a homomorphism $S(F) \to S(E)$ which we still denote by i. If i is injective on F then it is also injective on $S(F)$. In particular if $F \subset E$ and i is the inclusion mapping, we can identify $S(F)$ under i with the subalgebra of $S(E)$ generated by 1 and F.

We denote by $P_d(E)$ the space of all homogeneous polynomials on E of degree d. Similarly $S_d(E) = P_d(E')$. If $F \subset E$, it is clear that $S_d(F) = S(F) \cap S_d(E)$.

Now suppose E is the direct sum of two subspaces E_1, E_2. Let α_i denote the projection of E on E_i $(i = 1, 2)$ corresponding to the sum $E = E_1 + E_2$. Then given any polynomial function p on E_i, we can extend it to a polynomial function on E by setting $p(X) = p(\alpha_i X)$ $(X \in E)$. This defines an isomorphism of $P(E_i)$ into $P(E)$ and we can identify $P(E_i)$ with its image by means of it. In this way $P(E_1)$ and $P(E_2)$ become subalgebras of E and it is obvious that the mapping $(p_1, p_2) \rightarrow p_1 p_2$ $(p_i \in P(E_i)$ defines an isomorphism of $P(E_1) \otimes P(E_2)$ onto $P(E)$.

Let A be an automorphism of E. Then A can be extended uniquely to an automorphism $q \rightarrow q^A$ of $S(E)$. Moreover, for any $p \in P(E)$, define p^A to be the function $X \rightarrow p(A^{-1}X)$ $(X \in E)$. Then $p \rightarrow p^A$ is an automorphism of $P(E)$.

7. The radial component.[3] Assume k has characteristic zero and let $\mathfrak{g}$ be a Lie algebra over k. For any $X \in \mathfrak{g}$ let L_X denote the endomorphism of $S(\mathfrak{g})$ corresponding to multiplication by X in $S(\mathfrak{g})$. Also let d_X denote the unique derivation of $S(\mathfrak{g})$ which coincides on $\mathfrak{g}$ with $\operatorname{ad} X$. It is immediate that $d_X L_Y - L_Y d_X = L_{[X,Y]}$. For $Y \in \mathfrak{g}$, put[4]

$$\sigma_Y(X) = L_{[X,Y]} + d_X \qquad (X \in \mathfrak{g}).$$

One checks without difficulty that σ_Y is a representation of $\mathfrak{g}$ on $S(\mathfrak{g})$. Let $\mathfrak{G}$ be the universal enveloping algebra of $\mathfrak{g}$. We can extend σ_Y (uniquely) to a representation of $\mathfrak{G}$ which we again denote by σ_Y. Define a linear mapping Γ_Y of $\mathfrak{G} \otimes S(\mathfrak{g})$ into $S(\mathfrak{g})$, given by

$$\Gamma_Y(g \otimes p) = \sigma_Y(g)p \qquad (g \in \mathfrak{G}, p \in S(\mathfrak{g})).$$

Let λ denote the canonical mapping of $S(\mathfrak{g})$ onto $\mathfrak{G}$ (see [4(a), p. 192]). For any linear subspace U of $\mathfrak{g}$, define $\mathfrak{S}_d(U) = \lambda(S_d(U), \mathfrak{S}(U)) = \lambda(S(U))$ and $\mathfrak{S}_+(U) = \sum_{d \geqq 1} \mathfrak{S}_d(U)$. It is obvious that for a fixed $Y \in \mathfrak{g}$, Γ_Y defines a linear mapping of $\mathfrak{S}_{d_1}(\mathfrak{g}) \otimes S_{d_2}(\mathfrak{g})$ into $\sum_{0 \leqq d \leqq d_1 + d_2} S_d(\mathfrak{g})$.

Fix an element $X_0 \in \mathfrak{g}$ and let $\mathfrak{z}_{X_0}$ denote the centralizer of X_0 in $\mathfrak{g}$. Put $\mathfrak{g}_{X_0} = [X_0, \mathfrak{g}]$ and choose linear subspaces U and V in $\mathfrak{g}$ such that

$$\mathfrak{g} = U + \mathfrak{g}_{X_0} = V + \mathfrak{z}_{X_0}$$

---

[3] Cf. [4(d), § 5], [4(c), § 3] and [1, § 2].

[4] Cf. Lemma 3 of [4(c)].

where both sums are direct. For any $u \in U$, consider the linear mapping

$$\phi_u: (w, v) \to w + [X_0 + u, v] \qquad (w \in U, v \in V)$$

of $U \times V$ into $\mathfrak{g}$. It follows from the definition of U and V that ϕ_0 is bijective. Let $\mathcal{E}(\mathfrak{g})$ be the space of all endomorphisms of $\mathfrak{g}$ and I the identity mapping of $\mathfrak{g}$. Since

$$\phi_u(w, v) = \phi_0(w, v) + [u, v] \qquad (u, w \in U; v \in V),$$

it is clear that

$$u \to \phi_u\phi_0^{-1} - I$$

is a linear mapping of U into $\mathcal{E}(\mathfrak{g})$. Put

$$q(u) = \det(\phi_u\phi_0^{-1}) \qquad (u \in U).$$

Then $q \in P(U)$ and $q(0) = 1$.

LEMMA 8.[5] *Let u be an element in U such that $q(u) \neq 0$. Then Γ_{X_0+u} defines a bijective linear mapping of $\mathfrak{S}(V) \otimes S(U)$ onto $S(\mathfrak{g})$.*

We prove by induction on $d \geqq 0$ that

$$\sum_{d_1+d_2 \leqq d} \Gamma_{X_0+u}(\mathfrak{S}_{d_1}(V) \otimes S_{d_2}(U)) = \sum_{r \leqq d} S_r(\mathfrak{g}).$$

If $d = 0$, this is obvious. So assume $d \geqq 1$. Obviously the left side is contained in the right side. Hence in view of our induction hypothesis, it is enough to prove that

$$\sum_{d_1+d_2 \leqq d} \Gamma_{X_0+u}(\mathfrak{S}_{d_1}(V) \otimes S_{d_2}(U)) + \mathcal{S}_{d-1} \supset S_d(\mathfrak{g})$$

where $\mathcal{S}_r = \sum_{s \leqq r} S_s(\mathfrak{g})$. Now $q(u) \neq 0$ and therefore

$$\mathfrak{g} = \phi_u\phi_0^{-1}\mathfrak{g} = U + [X_0 + u, V].$$

Hence $S_d(\mathfrak{g})$ is spanned by elements α of the form

$$\alpha = \prod_{1 \leqq i \leqq d_1} w_i \prod_{1 \leqq j \leqq d_2} [X_0 + u, v_j]$$

where $w_i \in U$, $v_j \in V$ and $d_1 + d_2 = d$. Put $g = (-1)^{d_2}\lambda(v_1 v_2 \cdots v_{d_2}) \in \mathfrak{S}_{d_2}(V)$ and $p = w_1 w_2 \cdots w_{d_1} \in S_{d_1}(U)$. Then

$$\Gamma_{X_0+u}(g \otimes p) = \sigma_{X_0+u}(g)p \equiv \alpha \bmod \mathcal{S}_{d-1}.$$

This proves our assertion.

---

[5] Cf. Lemma 4 of [4(c)].

Now since ϕ_0 is bijective, $\dim \mathfrak{g} = \dim U + \dim V$. Therefore the spaces $\mathcal{S}_d$ and $\sum_{d_1+d_2\leqq d} \mathfrak{S}_{d_1}(V) \otimes S_{d_2}(U)$ have the same dimension. This proves that Γ_{X_0+u} is injective.

Let A and B be two vector spaces over k and suppose $\dim A < \infty$. Then $p\colon A \to B$ is called a polynomial mapping if the subspace B_1 of B spanned by $p(A)$, has finite dimension and p, viewed as a mapping of A into B_1, is a polynomial mapping in the usual sense.

LEMMA 9.[6] *Fix an element $p \in S(\mathfrak{g})$. Then there exists an integer $r \geqq 0$ and a polynomial mapping $\gamma_p\colon U \to \mathfrak{S}(V) \otimes S(U)$ such that*

$$\Gamma_{X_0+u}(\gamma_p(u)) = q(u)^r p$$

for $u \in U$.

Fix an integer $d \geqq 0$ and let

$$W_d = \sum_{d_1+d_2\leqq d} \mathfrak{S}_{d_1}(V) \otimes S_{d_2}(U)$$

and let $\Gamma(u)$ denote the restriction Γ_{X_0+u} on W_d. We have seen that $\dim W_d = \dim \mathcal{S}_d$. Fix a bijective linear mapping C of $\mathcal{S}_d$ onto W_d and put $A(u) = C\Gamma(u)$. Then it is obvious that $u \to A(u)$ is a polynomial mapping of U into the space $\mathcal{E}(W_d)$ of all endomorphisms of W_d. Let $Q(u) = \det A(u)$. Then it follows from the proof of Lemma 8 that $Q(u) = 0$ implies $q(u) = 0$ $(u \in U)$. Let $\bar{k}$ be the algebraic closure of k, $\bar{\mathfrak{g}} = \mathfrak{g} \otimes_k \bar{k}$ and $\bar{U}$, $\bar{V}$ the subspaces of $\bar{\mathfrak{g}}$ spanned by U and V respectively. We can regard Q and q as polynomial functions on $\bar{U}$. By applying the above reasoning to $\bar{\mathfrak{g}}$, $\bar{U}$, $\bar{V}$ instead of $\mathfrak{g}$, U, V, we deduce that $Q(\bar{u}) \neq 0$ unless $q(\bar{u}) = 0$ $(\bar{u} \in \bar{U})$. Since $\bar{k}$ is algebraically closed, this implies that Q divides q^r in $P(\bar{U})$ (and therefore also in $P(U)$) for some integer $r \geqq 0$.

Now let $N = \dim W_d$ and t an indeterminate. Then

$$\det(t - A(u)) = (-1)^N Q(u) + t \sum_{1\leqq j\leqq N} Q_j(u) t^{j-1} \qquad (u \in U)$$

where $Q_j \in P(U)$. Put

$$B(u) = (-1)^{N+1} \sum_{1\leqq j\leqq N} Q_j(u) A(u)^{j-1}.$$

Then B is a polynomial mapping of U into $\mathcal{E}(W_d)$ and

$$A(u)B(u) = B(u)A(u) = Q(u)I$$

---

[6] Cf. Lemma 5 of [4(c)].

where I is the identity mapping of W_d.

We can assume that d is so large that $p \in \mathcal{S}_d$. Then

$$\Gamma_{X_0+u}(B(u)Cp) = C^{-1}A(u)B(u)Cp = Q(u)p.$$

But $q^r = Qq_1$ $(q_1 \in P(U))$, as we saw above. Put

$$\gamma_p(u) = q_1(u)B(u)Cp.$$

Then

$$\Gamma_{X_0+u}(\gamma_p(u)) = q(u)^r p \qquad (u \in U).$$

Let U' be the set of all points $u \in U$ where $q(u) \neq 0$. Then $0 \in U'$ since $q(0) = 1$. For any $p \in S(\mathfrak{g})$, let $\beta_p(u)$ denote the unique element in $\mathfrak{S}(V) \otimes S(U)$ such that $p = \Gamma_{X_0+u}(\beta_p(u))$ $(u \in U')$. Moreover let $\alpha_u(p)$ denote the unique element in $S(U)$ such that

$$\beta_p(u) - (1 \otimes \alpha_u(p)) \in \mathfrak{S}_+(V) \otimes S(U) \qquad (u \in U').$$

We shall call $\alpha_u(p)$ the radical component[8] of p at u (with respect to X_0, U, V).

Remark. It is obvious from the proof of Lemma 8 that if $p \in S_d(\mathfrak{g})$ then $\alpha_u(p) \in \sum_{0 \leq r \leq d} S_r(U)$ for $u \in U'$.

8. The mapping Δ . Let $\mathfrak{g}$ be a Lie algebra over $\boldsymbol{R}$ and fix an element X in $\mathfrak{g}$. Select U and V as in § 7 so that $\mathfrak{g} = U + \mathfrak{g}_{X_0} = V + \mathfrak{z}_{X_0}$. Let $\mathfrak{g}_c$ denote the complexification of $\mathfrak{g}$ and U_c, V_c the subspaces of $\mathfrak{g}_c$ spanned by U, V respectively. Now we take $k = \boldsymbol{C}$ in § 7 and denote by $\mathfrak{G}$ the universal enveloping algebra of $\mathfrak{g}_c$. As before, U' is the set of all points $u \in U$ where $q(u) \neq 0$ (see Lemma 8).

$\mathfrak{g}$, being a real Euclidean space, is a manifold. For any $p \in S(\mathfrak{g}_c)$, we define, as usual (see [4(c), § 2] the differential operator $\partial(p)$ on $\mathfrak{g}$. Let D be a differential operator on an open subset Ω of $\mathfrak{g}$. Then we define a differential operator $\Delta(D)$ on the open subset $U' \cap (\Omega - X_0)$ of the Euclidean space U as follows. As usual let D_Y denote the local expression of D at $Y \in \Omega$ (see [4(c), p. 90]). Fix $u \in U' \cap (\Omega - X_0)$ and choose $p_u \in S(\mathfrak{g}_c)$ such that $D_{X_0+u} = \partial(p_u)$. Then

$$\Delta_u(D) = \partial(\alpha_u(p_u)).$$

Here $\Delta(D)$ is the local expression of $\Delta(D)$ at u and $\alpha_u(p_u)$ is the radial component of p_u at u (see § 7). Lemma 9 ensures that this does, indeed, define a differential operator on $U' \cap (\Omega - X_0)$.

For any $X \in \mathfrak{g}$, define a vector field $\tau(X)$ (cf. § 5) on $\mathfrak{g}$ by

$$\begin{aligned} f(Y;\tau(X)) &= -\{df(\exp(t\,\mathrm{ad}\,X)Y)/dt\}_{t=0} \\ &= -f(Y;\partial([X,Y])) \quad (f \in C^\infty(\mathfrak{g}), Y \in \mathfrak{g}). \end{aligned}$$

It is easy to verify that $[\tau(X),\tau(Y)] = \tau([X,Y])$ for X, $Y \in \mathfrak{g}$. Hence τ can be extended uniquely to a homomorphism of $\mathfrak{G}$ into the algebra of all differential operators on $\mathfrak{g}$.

A C^∞ function f on an open subset Ω of $\mathfrak{g}$ is called locally invariant if $\tau(X)f = 0$ for all $X \in \mathfrak{g}$. The significance of $\Delta(D)$ is given by the following result (cf. [4(c), Lemma 6]).

LEMMA 10. *Fix a point $u_0 \in U'$ and let f be a locally invariant C^∞ function on an open neighborhood Ω of $X_0 + u_0$ in $\mathfrak{g}$. Then if D is any differential operator on Ω,*

$$f(X_0 + u_0; D) = f(X_0 + u_0; \Delta(D)).$$

Choose $p \in S(\mathfrak{g}_c)$ such that $D_{X_0+u_0} = \partial(p)$. Then $\Delta_{u_0}(D) = \partial(\alpha_{u_0}(p))$. Hence it is enough to prove that

$$f(X_0 + u_0; \partial(p)) = f(X_0 + u_0; \partial(\alpha_{u_0}(p)))$$

for $p \in S(\mathfrak{g}_c)$.

Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then G operates on $\mathfrak{g}$ as follows: $X^x = \mathrm{Ad}(x)X$ $(x \in G, X \in \mathfrak{g})$. We sometimes also write xX for X^x. As usual (see [4(d), § 4]), we regard elements of $\mathfrak{G}$ as left-invariant differential operators on G. Select open neighborhoods G_0 and Ω_0 of 1 and $X_0 + u_0$ in G and Ω respectively, such that $\Omega_0{}^{G_0} \subset \Omega$. Put $F(x\colon X) = F(xX)$ $(x \in G_0, X \in \Omega_0)$ for any $F \in C^\infty(\Omega)$. Then we know [4(c), Lemma 3] that

$$F(x;g\colon X;\partial(p)) = F(x\colon X;\partial(\sigma_X(g)p)) = F(x\colon X;\partial(\Gamma_X(g \otimes p)))$$

for $g \in \mathfrak{G}$, $p \in S(\mathfrak{g}_c)$.

The mapping $X \to -X$ of $\mathfrak{g}_c$ can be extended uniquely to an anti-automorphism of $\mathfrak{G}$ which we denote by $*$.

LEMMA 11. *Let $F \in C^\infty(\Omega)$. Then*

$$F(x;g\colon X) = F(x\colon X;\tau(g^*))$$

for $x \in G_0$, $X \in \Omega_0$ and $g \in \mathfrak{G}$.

Let $Y_1, \cdots, Y_r \in \mathfrak{g}$. Then

$$\begin{aligned} &\{\partial^r F(x \exp t_1Y_1 \cdots \exp t_rY_r\colon X)/\partial t_1 \cdots \partial t_r\}_{t_1=\cdots=t_r=0} \\ &\qquad = \{\partial^r F(x\colon (\exp t_1Y_1 \cdots \exp t_rY_r)X)/\partial t_1 \cdots \partial t_r\}_{t_1=\cdots=t_r=0} \end{aligned}$$

for $x \in G_0$ and $X \in \Omega_0$. Lemma 11 follows immediately from this fact.

Now we return to Lemma 10. Fix $p \in S(\mathfrak{g}_0)$. Then, from Lemma 9, we can choose $g_i \in \mathfrak{S}_+(V_0)$, $p_j \in S(U_0)$, $q_{ij} \in P(U_0)$ $(1 \leqq i, j \leqq N)$ and an integer $r \geqq 0$ such that

$$p = \alpha_u(p) + q(u)^{-r} \sum_{1 \leqq i, j \leqq N} q_{ij}(u) \Gamma_{X_0+u}(g_i \otimes p_j)$$

for $u \in U'$. Therefore

$$\begin{aligned} f(X_0 + u; \partial(p)) = {} & f(X_0 + u; \partial(\alpha_u(p))) \\ & + q(u)^{-r} \sum_{i,j} q_{ij}(u) f(1; g_i \colon X_0 + u; \partial(p_j)) \end{aligned}$$

for $u \in U' \cap (\Omega - X_0)$. But g_i and therefore also g_i^*, are in $\mathfrak{S}_+(V_0)$. Hence

$$f(1; g_i \colon X_0 + u) = f(X_0 + u; \tau(g_i^*)) = 0$$

as f is locally invariant. Therefore

$$f(1; g_i \colon X_0 + u; \partial(p_j)) = 0 \qquad (1 \leqq i, j \leqq N)$$

and so

$$f(X_0 + u; \partial(p)) = f(X_0 + u; \partial(\alpha_u(p)))$$

for $u \in U' \cap (\Omega - X_0)$. This proves Lemma 10.

9. Application to invariant distributions.

From now on we shall assume that G is unimodular. Then $\mathfrak{g}$ is also unimodular, i. e. $\operatorname{tr} \operatorname{ad} X = 0$ for $X \in \mathfrak{g}$. Let dX denote the Euclidean measure on $\mathfrak{g}$ and dx the Haar measure on G (both normalized in some fixed but arbitrary way). Then dx is both left- and right-invariant and dX is invariant under G. Fix an orientation of $\mathfrak{g}$ and let $\omega_{\mathfrak{g}}$ be the positive differential form on $\mathfrak{g}$ of degree $n = \dim \mathfrak{g}$, such that[2] $\omega_{\mathfrak{g}} \sim dX$. Similarly let $\omega_G \neq 0$ be a real and left-invariant differential form on G of degree n. We can orient G in such a way that $\omega_G > 0$. Moreover we may assume that $\omega_G \sim dx$.

Now take $M = \mathfrak{g}$ and $\omega = \omega_{\mathfrak{g}}$ in the set up of § 5. Then the definitions of $\tau(X)$ $(X \in \mathfrak{g})$ of Sections 5 and 8 actually coincide and we conclude from Lemma 6 that $\tau(g^*) = \tau(g)^*$ $(g \in \mathfrak{G})$. Moreover g^* is the adjoint of the differential operator g on G, with respect to ω_G.

LEMMA 12. *Let π denote the mapping $(x, u) \to x(X_0 + u)$ of $G \times U$ into $\mathfrak{g}$. Then* $\operatorname{rank} d\pi = n$ *everywhere on* $G \times U'$.

Fix a point (x, u) in $G \times U'$. Since the tangent space at any point of a Euclidean space E can be canonically identified with E, the tangent space of $G\times U$ at (x,u) is $\mathfrak{g}\times U$. Similarly the tangent space of $\mathfrak{g}$ at $Y=x(X_0+u)$ is $\mathfrak{g}$. Let $f\in C^\infty(\mathfrak{g})$. Then

$$\{(d/dt)f(\pi(x\ \exp tX,u))\}_{t=0}=f(Y;\partial([X,X_0+u]^x))$$

for $X\in\mathfrak{g}$ and

$$\{(d/dt)f(\pi(x,u+tw))\}_{t=0}=f(Y;\partial(w^x))$$

for $w\in U$. Hence

$$(d\pi)_{x,u}(X,w)=([X,X_0+u]+w)^x$$

and therefore if $v\in V$, $w\in U$, it is clear that

$$(d\pi)_{x,u}(v,w)=(\phi_u(w,-v))^x$$

in the notation of § 7. But since $u\in U'$, ϕ_u is a bijective mapping of $U\times V$ onto $\mathfrak{g}$. This proves that $(d\pi)_{x,u}$ maps $V\times U$ onto $\mathfrak{g}^x=\mathfrak{g}$ and therefore rank $(d\pi)_{x,u}=n$.

Fix an orientation of U and let $\omega_U>0$ denote a differential form on U such that $\omega_U\sim du$, where du is a Euclidean measure on U. Now take $M=G\times U'$, $N=\pi(G\times U')$, $\omega_M\sim dx\,du$, $\omega_N\sim dX$. Then the following result is an immediate consequence of Theorem 1 and Lemma 12.

LEMMA 13. *Put* $\Omega=\pi(G\times U')=(X_0+U')^G$. *Then* Ω *is open in* $\mathfrak{g}$ *and for every* $\alpha\in C_c^\infty(G\times U')$, *there exists a unique element* $f_\alpha\in C_c^\infty(\Omega)$ *such that*

$$\int_{G\times U'}F(x(X_0+u))\alpha(x:u)\,dx\,du=\int_\Omega F(X)f_\alpha(X)\,dX$$

for every $F\in C_c^\infty(\Omega)$. *The mapping* $\alpha\to f_\alpha$ *of* $C_c^\infty(G\times U')$ *into* $C_c^\infty(\Omega)$ *is continuous and surjective and* $\operatorname{Supp} f_\alpha\subset\pi(\operatorname{Supp}\alpha)$.

COROLLARY 1. *Let* F *be a locally summable function on* Ω. *Then the function* $(x,u)\to F(x(X_0+u))$ *is locally summable on* $G\times U'$ *and*

$$\int_{G\times U'}F(x(X_0+u))\alpha(x:u)\,dx\,du=\int_\Omega F(X)f_\alpha(X)\,dX$$

for every $\alpha\in C_c^\infty(G\times U')$.

This follows from Corollary 2 of Theorem 1.

COROLLARY 2. *Let* Ω_0 *be an open subset of* Ω *and* T *a distribution on* Ω_0. *Select open subsets* G_0, U_0 *of* G *and* U' *respectively such that* $\pi(G_0\times U_0)\subset\Omega_0$. *Then the mapping*

$$\sigma_T' : \alpha \to T(f_\alpha) \qquad (\alpha \in C_c^\infty(G_0 \times U_0))$$

is a distribution on $G_0 \times U_0$. *Moreover* $\sigma_T' = 0$ *implies that* $T = 0$ *on* $\pi(G_0 \times U_0)$.

This follows by taking $M = G_0 \times U_0$ and $N = \pi(G_0 \times U_0)$ in Lemma 5.

For any $X \in \mathfrak{g}$, define the vector field X' on G as in Lemma 7. We denote the differential operator X' on G as in Lemma 7. We denote the differential operator $X' \times 1$ on $G \times U$ also by X'.

LEMMA 14. $f_{X'\alpha} = \tau(X) f_\alpha$ *for* $X \in \mathfrak{g}$ *and* $\alpha \in C_c^\infty(G \times U')$.

Fix α. Then for any $F \in C_c^\infty(\Omega)$,

$$\begin{aligned} \int F(x(X_0+u))\alpha(yx:u)\,dx\,du & \\ &= \int F(y^{-1}x(X_0+u))\alpha(x:u)\,dx\,du \\ &= \int F^y(x(X_0+u))\alpha(x:u)\,dx\,du \\ &= \int F^y(X) f_\alpha(X)\,dX \\ &= \int F(X) f_\alpha{}^{y^{-1}}(X)\,dX \end{aligned}$$

if y is sufficiently near 1 in G. Now for a given $Y \in \mathfrak{g}$, put $y_t = \exp tY$ $(t \in \boldsymbol{R})$. Then it is clear that

$$\begin{aligned} -\int F(x(X_0+u))\alpha(x;Y':u)\,dx\,du & \\ &= \{(d/dt)\int F(x(X_0+u))\alpha(y_t x:u)\,dx\,du\}_{t=0} \\ &= \{(d/dt)\int F(X) f_\alpha{}^{y_t^{-1}}(X)\,dX\}_{t=0} \\ &= -\int F(X) f_\alpha(X;\tau(Y))\,dX. \end{aligned}$$

This proves that $f_{Y'\alpha} = \tau(Y) f_\alpha$.

THEOREM 2. *Let* G_0 *and* U_0 *be open subsets of* G *and* U' *respectively. Assume that* G_0 *is connected and not empty. Then if* T *is a locally invariant distribution on* $\Omega_0 = \pi(G_0 \times U_0)$, *there exists a unique distribution* σ_T *on* U_0 *such that*

$$T(f_\alpha) = \sigma_T(\beta_\alpha) \qquad (\alpha \in C_c^\infty(G_0 \times U_0))$$

where β_α *is the function in* $C_c^\infty(U_0)$ *given by*

$$\beta_\alpha(u) = \int \alpha(x:u)\,dx \qquad (u \in U_0).$$

Moreover $\sigma_T = 0$ *implies that* $T = 0$.

Define the distribution σ_T' on $G_0 \times U_0$ as in Corollary 2 of Lemma 13. Fix $X \in \mathfrak{g}$. Then, as we have seen above, X' is a vector-field on $G \times U$.

We claim $X'\sigma_T' = 0$. Since dx is left-invariant, $(X')^* = -X'$ and therefore we have to prove that $\sigma_T'(X'\alpha) = 0$ for $\alpha \in C_c^\infty(G_0 \times U_0)$. Fix α. Then

$$\sigma_T'(X'\alpha) = T(f_{X'\alpha}) = T(\tau(X)f_\alpha) = 0$$

from Lemma 6, since T is locally invariant.

Fix an element $\beta \in C_c^\infty(U_0)$ and consider the mapping

$$S_\beta\colon \gamma \to \sigma_T'(\gamma \times \beta) \qquad (\gamma \in C_c^\infty(G_0)).$$

Then S_β is a distribution on G_0 and it is obvious from the above result that $X'S_\beta = 0$ $(X \in \mathfrak{g})$. Therefore from Lemma 7, there exists a complex number $C(\beta)$ such that

$$S_\beta(\gamma) = C(\beta)\int \gamma\, dx \qquad (\gamma \in C_c^\infty(G)).$$

Now fix $\gamma_0 \in C_c^\infty(G_0)$ such that $\int \gamma_0\, dx = 1$. Then

$$C(\beta) = S_\beta(\gamma_0) = \sigma_T'(\gamma_0 \times \beta).$$

The mapping $\beta \to \sigma_T'(\gamma_0 \times \beta)$ $(\beta \in C_c^\infty(U_0))$ is obviously a distribution on U_0, which we call σ_T. Then

$$\sigma_T'(\gamma \times \beta) = \sigma_T(\beta)\int \gamma\, dx \qquad (\gamma \in C_c^\infty(G_0), \beta \in C_c^\infty(U_0)).$$

Now for any $\alpha \in C_c^\infty(G_0 \times U_0)$ define β_α as in the statement of Theorem 2. Then the mapping $\sigma\colon \alpha \to \sigma_T(\beta_\alpha)$ is obviously a distribution on $G_0 \times U_0$ and $\sigma_T'(\gamma \times \beta) = \sigma(\gamma \times \beta)$ for $\beta \in C_c^\infty(U_0)$, $\gamma \in C_c^\infty(G_0)$. Hence we conclude from Lemma 3 that $\sigma_T' = \sigma$. This proves that

$$T(f_\alpha) = \sigma_T'(\alpha) = \sigma_T(\beta_\alpha) \qquad (\alpha \in C_c^\infty(G_0 \times U_0).$$

Moreover for a given $\beta \in C_c^\infty(U_0)$, set $\alpha = \gamma_0 \times \beta$. Then $\beta = \beta_\alpha$. This shows that the mapping $\alpha \to \beta_\alpha$ of $C_c^\infty(G_0 \times U_0)$ into $C_c^\infty(U_0)$ is surjective. Hence the uniqueness of σ_T in Theorem 2 is obvious. Finally suppose $\sigma_T = 0$. Then $\sigma_T'(\alpha) = \sigma_T(\beta_\alpha) = 0$ $(\alpha \in C_c^\infty, G_0 \times U_0))$ and therefore $\sigma_T' = 0$. But, as we have seen (Corollary 2 of Lemma 13), this implies that $T = 0$.

We keep to the notation of Theorem 2.

THEOREM 3.[7] *Assume further that* $1 \in G_0$. *Then if* D *is a differential operator and* T *a distribution on* Ω_0 *and they are both locally invariant, we have* $\sigma_{DT} = \Delta(D)\sigma_T$.

We observe that DT is also locally invariant and therefore σ_{DT} is defined. Moreover $U_0 \subset \Omega_0 - X_0$ and therefore $\Delta(D)$ is defined on U_0. Theorems 2 and 3 provide the main justification for the consideration of the differential operator $\Delta(D)$.

---

[7] This can be considered as a generalization of Lemma 10 to distributions.

COROLLARY. *Let D_1, D_2 be two differential operators and T a distribution on Ω_0. Suppose all three of them are locally invariant. Then*

$$\Delta(D_1D_2)\sigma_T = \Delta(D_1)\Delta(D_2)\sigma_T.$$

This an immediate consequence of Theorem 3.

The proof of Theorem 3 requires some preparation.

LEMMA 15. *Let D be a locally invariant differential operator on an open set Ω in $\mathfrak{g}$. Fix an element $X \in \Omega$ and an open connected neighborhood G_0 of 1 in G such that $yX \in \Omega$ for all $y \in G_0$. Then*

$$D_{yX} = (D_X)^y \qquad (y \in G_0).$$

For any $x \in G$, let $p \to p^x$ $(p \in S(\mathfrak{g}_c)$ denote the unique automorphism of $S(\mathfrak{g}_c)$ which coincides with $\mathrm{Ad}(x)$ on $\mathfrak{g}_c$. In this way G operates on $S(\mathfrak{g}_c)$ and it is obvious that for any $d \geqq 0$, the space $S_d(\mathfrak{g}_c)$ is stable under G. Hence for a given p in $S(\mathfrak{g}_c)$, the elements p^x $(x \in G)$ span a finite-dimensional space V. The corresponding representation on V being obviously continuous, it follows, from the theory of Lie groups, that it is analytic. It is easy to verify that $\{dp^{\exp tY}/dt\}_{t=0} = d_Yp$ $(Y \in \mathfrak{g})$. Moreover it is clear that $\partial(Y^x) = (\partial(Y))^x$ $(Y \in \mathfrak{g}, x \in G)$ and therefore $\partial(p^x) = (\partial(p))^x$ $(p \in S(\mathfrak{g}_c))$.

We can choose $p_i \in S(\mathfrak{g}_c)$ and $a_i \in C^\infty(\Omega)$ $(1 \leqq i \leqq r)$ such that

$$D = \sum_{1 \leqq i \leqq r} a_i\partial(p_i).$$

Moreover we may obviously assume that $p_1, \cdots, p_r$ are linearly independent over $\boldsymbol{C}$ and the subspace of $S(\mathfrak{g}_c)$ spanned by them is stable under G. Let $p_i^x = \sum_j p_jc_{ji}(x)$ $(1 \leqq i \leqq r, x \in G)$ where c_{ji} are analytic functions on G. Let $C(x)$ denote the matrix $(c_{ij}(x))_{1 \leqq i, j \leqq r}$. Then $x \to C(x)$ is a matrix representation of G. Let $Y \to C(Y)$ $(Y \in \mathfrak{g})$ denote the corresponding representation of $\mathfrak{g}$ and $c_{ij}(Y)$ $(1 \leqq i, j \leqq r)$ the coefficients of the matrix $C(Y)$.

Put $g_i(y) = \sum_j c_{ij}(y^{-1})a_j(yX)$ $(y \in G_0)$. Then if $Y \in \mathfrak{g}$,

$$\begin{aligned} -g_i(y; Y') &= \{dg_i(\exp tY \cdot y)/dt\}_{t=0} \\ &= -\sum_j c_{ij}(y^{-1})\{a_j(yX; \tau(Y)) + \sum_k c_{jk}(Y)a_k(yX)\}. \end{aligned}$$

We claim that if $a \in C^\infty(\Omega)$ and $p \in S(\mathfrak{g}_c)$,

$$\begin{aligned} \tau(Y) \circ a - a \circ \tau(Y) &= \tau(Y)a, \\ \tau(Y) \circ \partial(p) - \partial(p) \circ \tau(Y) &= \partial(d_Yp). \end{aligned}$$

The first assertion is obvious since $\tau(Y)$ is a vector field. The second is seen as follows. Let $f \in C^\infty(\mathfrak{g})$. Then

$$\tau(Y)(\partial(p)f) = \{d(\partial(p)f)^{y_t}/dt\}_{t=0}$$

where $y_t = \exp tY$. But $(\partial(p)f)^{y_t} = \partial(p^{y_t})f^{y_t}$ and therefore

$$\tau(Y)(\partial(p)f) = \partial(d_Y p)f + \partial(p)\,(\tau(Y)f).$$

Now since D is locally invariant, we have

$$0 = \tau(Y)\circ D - D\circ\tau(Y) = \sum_i\{(\tau(Y)a_i)\partial(p_i) + a_i\partial(d_Y p_i)\}.$$

But $d_Y p_i = \sum_j p_j c_{ji}(Y)$. Hence

$$\sum_i\{\tau(Y)a_i + \sum_j c_{ij}(Y)a_j\}\partial(p_i) = 0.$$

From this it is clear that

$$\tau(Y)a_i + \sum_j c_{ij}(Y)a_j = 0$$

But this obviously implies that $g_i(y;Y') = 0$. Since G_0 is connected, this shows that g_i is a constant. Hence

$$g_i(y) = g_i(1) = a_i(X).$$

Therefore

$$a_i(X) = \sum_j c_{ij}(y^{-1})a_j(yX)$$

or

$$a_i(yX) = \sum_j c_{ij}(y)a_j(X) \qquad (y \in G).$$

But then

$$(D_X)^y = \sum_i a_i(X)\partial(p_i^y) = \sum_{j,i} c_{ji}(y)a_i(X)\partial(p_j)$$
$$= \sum_j a_j(yX)\partial(p_j) = D_{yX}.$$

This proves Lemma 15.

Now we come to the proof of Theorem 3. Fix $\alpha \in C_c^\infty(G_0 \times U_0)$ and $F \in C_c^\infty(\Omega_0)$. Then

$$\int FD^*f_\alpha\, dX = \int DF\cdot f_\alpha\, dX = \int F(x(X_0+u);D)\alpha(x:u)\,dx\,du$$

from the definition of f_α. But

$$F(x(X_0+u);D) = F(x(X_0+u);D_{x(X_0+u)})$$
$$= F(x(X_0+u);(D_{X_0+u})^x) \quad (x \in G_0, u \in U_0)$$

from Lemma 15. However

$$F(x(X_0+u)\,;(D_{X_0+u})^{x})=F(x\colon X_0+u;D_{X_0+u})$$

in the notation of Lemma 11. Moreover we know (see the proof of Lemma 10) that there exist elements $g_i\in\mathfrak{S}_+(V_0)$, $p_i\in S(U_0)$ and $a_i\in C^\infty(U_0)$ $(1\leqq i\leqq r)$ such that

$$D_{X_0+u}=\Delta_u(D)+\sum_{1\leqq i\leqq r}a_i(u)\partial(\Gamma_{X_0+u}(g_i\otimes p_i)) \qquad (u\in U_0).$$

Hence

$$\begin{aligned}&F(x\colon X_0+u;D_{X_0+u})\\&\quad=F(x\colon X_0+u;\Delta(D))+\sum_i a_i(u)F(x;g_i\colon X_0+u;\partial(p_i))\end{aligned}$$

for $x\in G_0$ and $u\in U_0$. Put

$$\begin{aligned}\alpha_i(x\colon u)&=\alpha(x;g_i{}^*\colon u;\partial(p_i)^*\circ a_i), &&(1\leqq i\leqq r)\\ \alpha_0(x\colon u)&=\alpha(x\colon u;\Delta(D)^*),\end{aligned}$$

where, as usual, * denotes adjoint. Then

$$\begin{aligned}\int FD^*f_\alpha\,dX&=\sum_{0\leqq i\leqq r}\int F(x(X_0+u))\alpha_i(x\colon u)\,dx\,du\\&=\sum_{0\leqq i\leqq r}\int Ff_{\alpha_i}\,dX.\end{aligned}$$

This shows that

$$D^*f_\alpha=\sum_{0\leqq i\leqq r}f_{\alpha_i}$$

and therefore

$$T(D^*f_\alpha)=\sum_{0\leqq i\leqq r}\sigma_T(\beta_{\alpha_i})$$

from Theorem 2. However g_i and therefore also $g_i{}^*$, are in $\mathfrak{S}_+(V_0)$. Hence

$$\int\alpha(x;g_i{}^*\colon u)\,dx=0 \qquad (u\in U_0)$$

from the right-invariance of the measure dx. This shows that $\beta_{\alpha_i}=0$ for $i\geqq 1$. Moreover it is obvious that $\beta_{\alpha_0}=\Delta(D)^*\beta_\alpha$. Therefore

$$T(D^*f_\alpha)=\sigma_T(\Delta(D)^*\beta_\alpha)$$

and so we can conclude from Theorem 2 that $\sigma_{DT}=\Delta(D)\sigma_T$. This completes the proof of Theorem 3.

10. Statement of Theorem 4. Now we intend to apply the results of § 9 to prove a theorem about distributions on a semisimple Lie algebra. So let us assume that $\mathfrak{g}$ is semisimple and G is connected. Identify $\mathfrak{g}_0$ with its dual, by means of the bilinear form $B(X,Y)=\mathrm{tr}(\mathrm{ad}\,X\,\mathrm{ad}\,Y)$ $(X,Y\in\mathfrak{g}_0)$. In view of the invariance of B under G, this identification is compatible with the action of G. Thus elements of $S(\mathfrak{g}_c)$ become polynomial functions on $\mathfrak{g}_c$. Let $I(\mathfrak{g}_c)$ denote the subalgebra of all invariants of G in $S(\mathfrak{g}_c)$. It is easy to see that $I(\mathfrak{g}_c)$ is the set of all $p \in S(\mathfrak{g}_c)$ such that $d_X p = 0$ for all $X \in \mathfrak{g}$. By the Killing form (of $\mathfrak{g}$), we mean the polynomial function $\omega : X \to \operatorname{tr}(\operatorname{ad} X)^2 (X \in \mathfrak{g}_c)$. If θ is any automorphism of $\mathfrak{g}_c$, it is clear that $\operatorname{ad}(\theta X) = \theta(\operatorname{ad} X)\theta^{-1}$ and therefore $\omega(\theta X) = \omega(X)\, (X \in \mathfrak{g}_c)$. Hence, in particular, $\omega \in I(\mathfrak{g}_c)$.

An element $X \in \mathfrak{g}_c$ is called nilpotent if $\operatorname{ad} X$ is nilpotent. Let $\mathfrak{N}$ denote the set of all nilpotent elements of $\mathfrak{g}$.

THEOREM 4. *Let T be a distribution on an open subset Ω of $\mathfrak{g}$. Assume that*

1) *T is locally invariant,*

2) $\operatorname{Supp} T \subset \mathfrak{N} \cap \Omega$,

3) *$\partial(\omega) T = 0$, where ω is the Killing form of $\mathfrak{g}$.*

Then $T = 0$.

The proof of this theorem is rather long and requires considerable preparation. It is obvious that $\mathfrak{N}^x = \mathfrak{N}\, (x \in G)$. Hence G operates on $\mathfrak{N}$.

LEMMA 16. *$\mathfrak{N}$ is the union of a finite number of G-orbits.*

This follows from a result of Kostant (Corollary 3.7 and Lemma 5.1 of [6]) and Proposition 2.3 of [2].

Let $\mathfrak{N}_q$ denote the union of all G-orbits in $\mathfrak{N}$ of topological dimension $\leqq q$. Also put $\mathfrak{N}_{-1} = \emptyset$ and let ${}^c\mathfrak{N}_q$ denote the complement of $\mathfrak{N}_q$ in $\mathfrak{N}$.

LEMMA 17. *$\mathfrak{N}_q$ is closed in $\mathfrak{N}$ and if $X_0 \in \mathfrak{N}_q \cap {}^c\mathfrak{N}_{q-1}$, then X_0^G is open in $\mathfrak{N}_q$ $(q \geqq 0)$.*

This is an easy consequence of the theory of algebraic groups. We shall give a proof (which is due to Borel) in § 18.

COROLLARY. *For any $X_0 \in \mathfrak{N}$, the orbit X_0^G is locally compact.*

Let q be the topological dimension of X_0^G. Then X_0^G is open in the closed subset $\mathfrak{N}_q$ of $\mathfrak{N}$. Since $\mathfrak{N}$ is obviously closed in $\mathfrak{g}$, $\mathfrak{N}_q$ is locally compact and our assertion follows.

Fix $X_0 \in \mathfrak{N}$ and let $Z(X_0)$ denote the centralizer of X_0 in G. Then $Z(X_0)$ is a closed subgroup of G. Let $x \to x^*$ denote the natural mapping of G on $G^* = G/Z(X_0)$ and define $x^* X_0 = X_0^{x^*} = xX_0$.

LEMMA 18. *$x^* \to x^* X_0$ is a homeomorphism of G^* on X_0^G.*

This follows from a theorem of Arens (see [7, p. 65]), if we take the corollary of Lemma 17 into account.

Let $\mathfrak{z}_{X_0}$ denote, as before, the centralizer of X_0 in $\mathfrak{g}$.

COROLLARY. *Let q be the topological dimension of $X_0{}^G$. Then*

$$q = \dim G - \dim Z(X_0) = \dim \mathfrak{g} - \dim \mathfrak{z}_{X_0}.$$

This is obvious from Lemma 18.

We observe that in order to prove Theorem 4 it would be sufficient to obtain the following result.

LEMMA 19. *Fix an integer $q \geqq 0$ and let T be a locally invariant distribution on an open subset Ω of $\mathfrak{g}$ such that $\partial(\omega)T = 0$ and* $\operatorname{Supp} T \subset \mathfrak{N}_q \cap \Omega$. *Then* $\operatorname{Supp} T \subset \mathfrak{N}_{q-1} \cap \Omega$.

11. **Some elementary facts about distributions.** Let E be a finite-dimensional vector space over $\boldsymbol{R}$ and E_c its complexification. Fix a point $X_0 \in E$ and let δ_{X_0} denote the distribution $\delta_{X_0}: f \to f(X_0)$ $(f \in C_c{}^\infty(E))$ on E.

LEMMA 20. *Let σ be a distribution on an open neighborhood U of X_0 in E and suppose* $\operatorname{Supp} \sigma \subset \{X_0\}$. *Then there exists a unique $p \in S(E_c)$ such that*[8] $\sigma = \partial(p)\delta_{X_0}$.

This is a well-known and elementary result (see [9, Theorem 35, p. 99]).

LEMMA 21. *Let σ be a distribution on an open set U in E and X_0 a point in U. Let η be a function in $C^\infty(U)$ such that $\eta(X) = 0$ for every $X \in$* $\operatorname{Supp} \sigma$. *Then there exists an integer $k \geqq 0$ such that $X_0 \notin$* $\operatorname{Supp}(\eta^k\sigma)$.

Fix an open neighborhood V of X_0 in U such that V is relatively compact in U. Let $X_1, \cdots, X_n$ be a base for E over $\boldsymbol{R}$ and let F_m denote the set of all monomials $X_1{}^{m_1} \cdots X_n{}^{m_n} \in S(E_c)$ with $m_1 + \cdots + m_n \leqq m$. If D is any differential operator on U, we can choose $m \geqq 0$ and $a_p \in C^\infty(U)$ $(p \in F_m)$ such that

$$D = \sum_{p \in F_m} a_p \partial(p).$$

Hence if $f \in C_c{}^\infty(V)$, it is clear that

$$|Df| \leqq c \max_{p \in F_m} \sup |\partial(p)f|$$

where $c = \sum_{p \in F_m} \sup_V |a_p|$.

---

[8] Here the adjoint is defined with respect to the Euclidean measure on E.

Now, from Lemma 2, we can choose differential operators $D_1, \cdots, D_N$ on U such that

$$|\sigma(f)| \leqq \sum_{1 \leqq j \leqq N} \sup |D_j f| \qquad (f \in C_c^\infty(V)).$$

Hence it is obvious that we can choose an integer $m \geqq 0$ and a number $b \geqq 0$ such that

$$|\sigma(f)| \leqq b \max_{p \in F_m} \sup |\partial(p) f| \qquad (f \in C_c^\infty(V)).$$

First assume that η is a real-valued function. Fix $f \in C_c^\infty(V)$ and for any $\alpha \in C_c^\infty(\boldsymbol{R})$, define

$$\tau_f(\alpha) = \sigma(\alpha(\eta) f)$$

where $\alpha(\eta)$ denotes the function $X \rightarrow \alpha(\eta(X))$ on U. Then τ_f is obviously a distribution on $\boldsymbol{R}$. We claim $\operatorname{Supp} \tau_f \subset \{0\}$. For suppose $0 \notin \operatorname{Supp} \alpha$. Then if $X \in \operatorname{Supp}(\alpha(\eta) f)$, it is clear that $\eta(X) \in \operatorname{Supp} \alpha$ and $X \in V$. Hence $\eta(X) \neq 0$ and therefore, in view of our hypothesis, $X \notin \operatorname{Supp} \sigma$. This shows that $\operatorname{Supp}(\alpha(\eta) f) \cap \operatorname{Supp} \sigma = \emptyset$ and therefore $\tau_f(\alpha) = \sigma(\alpha(\eta) f) = 0$. Hence $\operatorname{Supp} \tau_f \subset \{0\}$.

On the other hand

$$|\tau_f(\alpha)| = |\sigma(\alpha(\eta) f)| \leqq b \max_{p \in F_m} \sup |\partial(p)(\alpha(\eta) f)|$$
$$\leqq c \max_{0 \leqq r \leqq m} \sup |d^r \alpha / dt^r|$$

where c is a positive number independent of α and t is the coordinate variable on $\boldsymbol{R}$. Therefore (see [9, pp. 99-100]) we can choose complex numbers $c_r(f)$ $(0 \leqq r \leqq m)$ such that

$$\tau_f(\alpha) = \sum_{0 \leqq r \leqq m} c_r(f) (d^r \alpha / dt^r)_0 \qquad (\alpha \in C_c^\infty(\boldsymbol{R})),$$

where $(d^r \alpha / dt^r)_0$ is the value of $d^r \alpha / dt^r$ at $t = 0$. Put $\alpha = t^{m+1} \beta$ for any $\beta \in C_c^\infty(\boldsymbol{R})$. Then $(d^r \alpha / dt^r)_0 = 0$ $(0 \leqq r \leqq m)$ and therefore

$$0 = \sigma(\alpha(\eta) f) = \sigma(\eta^{m+1} \beta(\eta) f).$$

Since V is relatively compact in U, we can choose a compact set J in $\boldsymbol{R}$ such that $\eta(V) \subset J$. Choose $\beta \in C_c^\infty(\boldsymbol{R})$ such that $\beta = 1$ on J. Then if $X \in V$, $\beta(\eta(X)) = 1$ and therefore $\beta(\eta) f = f$. Hence

$$\sigma(\eta^{m+1} f) = \sigma(\eta^{m+1} \beta(\eta) f) = 0.$$

This proves that $\eta^{m+1} \sigma = 0$ on V.

Now we come to the general case. Let $\eta = \eta_R + \sqrt{-1}\, \eta_I$ where η_R and

η_I are real-valued functions on U. Then by the above result, we can choose an integer $l \geqq 0$ such that $X_0 \notin \operatorname{Supp} \eta_R{}^l \sigma$, $X_0 \notin \operatorname{Supp} \eta_I{}^l \sigma$. Therefore

$$\eta^{2l}\sigma = (\eta_R + \sqrt{-1}\,\eta_I)^{2l}\sigma = 0$$

around X_0. This completes the proof.

12. First step in the proof of Theorem 4. We now return to the proof of Lemma 19. First assume $q = 0$. Then if $X_0 \in \mathfrak{N}_0$, it follows from the corollary of Lemma 18 that $\mathfrak{g} = \mathfrak{z}_{X_0}$. Since $\mathfrak{g}$ is semisimple, this implies that $X_0 = 0$. Therefore $\mathfrak{N}_0 = \{0\}$ and so if $0 \notin \Omega$, it is clear that $\operatorname{Supp} T = \emptyset$ and hence $T = 0$. So let us assume that $0 \in \Omega$. Then $T = \partial(p)\delta_0$ for some $p \in S(\mathfrak{g}_c)$, by Lemma 20. But $\partial(\omega)T = 0$ and therefore $\partial(\omega p)\delta_0 = 0$. This, however, implies, again by Lemma 20, that $\omega p = 0$. Since $S(\mathfrak{g}_c)$ is obviously an integral domain, we get $p = 0$ and therefore $T = \partial(p)\delta_0 = 0$.

So we can now assume that $q \geqq 1$. Fix an element $X_0 \in \mathfrak{N}_q \cap {}^c\mathfrak{N}_{q-1}$. Then we have to prove that $X_0 \notin \operatorname{Supp} T$. Choose U and V as in § 8 and use the notation of § 9. Then $0 \in U'$.

LEMMA 22. *We can choose an open neighborhood U_0 of zero in U' with the following property. Let x be an element in G. Then $xX_0 - X_0 \notin U_0$ unless $xX_0 = X_0$.*

I give here an improved version of my original proof, which was suggested by Kneser. It follows from Lemma 18 that $x^* \to x^*X_0$ is a univalent and regular analytic mapping of G^* into $\mathfrak{g}$. Hence $X_0{}^G$ can be regarded as an analytic submanifold of $\mathfrak{g}$ (see [3(a), p. 85]). Its tangent space at X_0 is obviously $\mathfrak{g}_{X_0}$. On the other hand $X_0 + U$ is an affine variety in $\mathfrak{g}$, whose tangent space at X_0 is U. Since $\mathfrak{g}_{X_0} \cap U = \{0\}$, $X_0{}^G$ and $X_0 + U$ are transversal at X_0. Therefore we can choose an open neighborhood $G_0{}^*$ of 1^* in G^* such that $G_0{}^*X_0 \cap (X_0 + U) = \{X_0\}$. On the other hand, by Lemma 18, there exists an open neighborhood $\mathfrak{g}_0$ of X_0 in $\mathfrak{g}$ such that $G_0{}^*X_0 = \mathfrak{g}_0 \cap X_0{}^G$. Now select an open neighborhood U_0 of zero in U' such that $X_0 + U_0 \subset \mathfrak{g}_0$. Then

$$X_0{}^G \cap (X_0 + U_0) \subset G_0{}^*X_0 \cap (X_0 + U) = \{X_0\}.$$

This proves the lemma.

In order to prove that $X_0 \notin \operatorname{Supp} T$, we may obviously assume that $X_0 \in \Omega$. By Lemma 17, $X_0{}^G$ is open in $\mathfrak{N}_q$. Hence we can choose an open neighborhood Ω_0 of X_0 in Ω such that $\Omega_0 \cap \mathfrak{N}_q \subset X_0{}^G$. Let G_0 be an open connected neighborhood of 1 in G. We can assume that G_0 and U_0 (as defined in

Lemma 22) are so small that $\pi(G_0 \times U_0) \subset \Omega_0$. (Here π has the same meaning as in § 9.) Define the distribution σ_T on U_0 as in Theorem 2.

LEMMA 23. Supp $\sigma_T \subset \{0\}$.

Fix $\beta \in C_c^\infty(U_0)$ such that $0 \notin \operatorname{Supp} \beta$. Then we have to show that $\sigma_T(\beta) = 0$. Choose $\gamma \in C_c^\infty(G_0)$ such that $\int \gamma \, dx = 1$ and put $\alpha = \gamma \times \beta$. Then $\sigma_T(\beta) = T(f_\alpha)$. But we know that $\operatorname{Supp} f_\alpha \subset \pi(\operatorname{Supp} \alpha)$. Hence if $X \in \operatorname{Supp} f_\alpha$, we have $X = x(X_0 + u)$ for some $x \in \operatorname{Supp} \gamma$ and $u \in \operatorname{Supp} \beta$. We claim $X \notin X_0{}^G$. For otherwise $X = yX_0$ for some $y \in G$. Therefore $X_0 + u = zX_0$ where $z = x^{-1}y$. Hence $u = zX_0 - X_0 \in U_0$. But $u \neq 0$ since $u \in \operatorname{Supp} \beta$. So we get a contradiction with the definition of U_0 (see Lemma 22). This shows that $\operatorname{Supp} f_\alpha \cap X_0{}^G = \emptyset$. However

$$\operatorname{Supp} T \cap \operatorname{Supp} f_\alpha \subset \mathfrak{N}_g \cap \Omega_0 \subset X_0{}^G.$$

Therefore $\operatorname{Supp} T \cap \operatorname{Supp} f_\alpha = \emptyset$ and so $T(f_\alpha) = 0$. This proves the lemma.

COROLLARY. *There exists a unique element* $z \in S(U_c)$ *such that*

$$\sigma_T(\beta) = \beta(0; \partial(z))$$

for all $\beta \in C_c^\infty(U_0)$.

This is an immediate consequence of Lemma 20.

Remark. In order to prove that $X \notin \operatorname{Supp} T$, it would, in view of Theorem 2, be enough to show that $\sigma_T = 0$. Hence it would be sufficient to establish that $z = 0$ in the above corollary.

13. Theorems of Jacobson-Morosow and Mostow.

In this section we recall some known facts about semisimple Lie algebras.

LEMMA 24 (Jacobson-Morosow). *Let* $\mathfrak{g}$ *be a semisimple Lie algebra over a field* k *of characteristic zero. If* $X \neq 0$ *is any nilpotent element in* $\mathfrak{g}$, *we can choose two elements* $H, Y \in \mathfrak{g}$ *such that*

$$[H, X] = 2X, \quad [H, Y] = -2Y, \quad [X, Y] = H.$$

For a proof see [5, p. 108, Theorem 3].

LEMMA 25. *Let* $\mathfrak{l}$ *be a three dimensional Lie algebra over* k *spanned by three elements* H, X, Y *with the relations*

$$[H, X] = 2X, \quad [H, Y] = -2Y, \quad [X, Y] = H.$$

Then $\mathfrak{l}$ *is semisimple. Let* ρ *be a representation of* $\mathfrak{l}$ *on a finite-dimensional* *vector space* V. *Then* 1) ρ *is fully reducible,* 2) $\rho(X)$, $\rho(Y)$ *are nilpotent,* 3) $\rho(H)$ *is semisimple and all its eigenvalues are rational integers. Now suppose* V *is irreducible and* $\dim V = m+1$ $(m \geqq 0)$. *Then the eigenvalues of* $\rho(H)$ *are* $m-2r$ $(0 \leqq r \leqq m)$ *and each of them has multiplicity* 1. *Let* V_m *be the subspace of* V *corresponding to the eigenvalue* m. *Then* V_m *is also the space of all* $v \in V$ *such that* $\rho(X)v = 0$. *Finally* ρ *is absolutely irreducible and, apart from equivalence, there is exactly one irreducible representation of* $\mathfrak{l}$ *of degree* d *for any integer* $d \geqq 1$.

All this is quite elementary and well known.

Let $\mathfrak{g}$ be a semisimple Lie algebra over $\boldsymbol{R}$. A Cartan involution θ of $\mathfrak{g}$ is an isomorphism of $\mathfrak{g}$ such that $\theta^2 = 1$ and the quadratic form $B(X,\theta(X)) = \operatorname{tr}(\operatorname{ad} X \operatorname{ad} \theta X)$ $(X \in \mathfrak{g})$ is negative-definite on $\mathfrak{g}$. The following result is due to Mostow [8, Theorem 6, p. 53].

LEMMA 26 (Mostow). *Let* $\mathfrak{g}$ *be a semisimple Lie algebra over* $\boldsymbol{R}$ *and* $\mathfrak{g}_1$ *a semisimple subalgebra of* $\mathfrak{g}$. *Let* θ_1 *be a Cartan involution of* $\mathfrak{g}_1$. *Then* θ_1 *can be extended to a Cartan involution* θ *of* $\mathfrak{g}$.

14. Some computations. We now return to the notation of § 12. Since $q = \dim \mathfrak{g} - \dim \mathfrak{z}_{X_0} \geqq 1$, $X_0 \neq 0$. Therefore by Lemma 24, we can choose $H_0, Y_0 \in \mathfrak{g}$ such that $[H_0, X_0] = 2X_0$, $[H_0, Y_0] = -2Y_0$, $[X_0, Y_0] = H_0$. Let $\mathfrak{l} = \boldsymbol{R}H_0 + \boldsymbol{R}X_0 + \boldsymbol{R}Y_0$. It is easy to see that $\mathfrak{l}$ is a three-dimensional subalgebra of $\mathfrak{g}$ and the linear mapping

$$\theta_0 : (X_0, Y_0, H_0) \to (-Y_0, -X_0, -H_0)$$

defines a Cartan involution of $\mathfrak{l}$. Therefore, from Lemma 26, we can extend it to a Cartan involution θ of $\mathfrak{g}$. Introduce the structure of a real Hilbert space on $\mathfrak{g}$ by means of the positive-definite quadratic form

$$\| Z \|^2 = -B(Z, \theta(Z)) \qquad (Z \in \mathfrak{g}).$$

LEMMA 27. *Let* $X \in \mathfrak{g}$. *Then the Hilbert-space adjoint of* $\operatorname{ad} X$ *is* $-\operatorname{ad}\theta(X)$.

Let $\langle Y, Z\rangle = -B(Y,\theta(Z)) = -B(\theta(Y), Z)$ $(Y, Z \in \mathfrak{g})$ denote the scalar product in the Hilbert space $\mathfrak{g}$. Then

$$\begin{aligned} \langle [X,Y], Z\rangle &= -B([\theta(X),\theta(Y)], Z) \\ &= B(\theta(Y), [\theta(X), Z]) \\ &= -\langle Y, [\theta(X), Z]\rangle \end{aligned}$$

and this is equivalent to our assertion.

We can take (in § 12) U and V to be the orthogonal complements of $\mathfrak{g}_{X_0}$ and $\mathfrak{z}_{X_0}$ respectively in $\mathfrak{g}$. Then we have the following result.

LEMMA 28. *$U=\theta(\mathfrak{z}_{X_0})$ and $V=\theta(\mathfrak{g}_{X_0})$.*

Let $Z\in\mathfrak{z}_{X_0}$ and $Y\in\mathfrak{g}$. Then

$$\langle\theta(Z),\ [X_0,Y]\rangle=-\langle[\theta(X_0),\theta(Z)],Y\rangle$$

from Lemma 27. But $[\theta(X_0),\theta(Z)]=\theta([X_0,Z])=0$. This proves that $\theta(\mathfrak{z}_{X_0})$ is orthogonal to $\mathfrak{g}_{X_0}$. But

$$\dim\theta(\mathfrak{z}_{X_0})=\dim\mathfrak{z}_{X_0}=\dim\mathfrak{g}-\dim\mathfrak{g}_{X_0}=\dim U.$$

Hence $\theta(\mathfrak{z}_{X_0})=U$.

Since B is invariant under θ, it is clear that θ is unitary. Therefore the orthogonal complement of $\theta(\mathfrak{g}_{X_0})$ is $\theta^2(\mathfrak{z}_{X_0})=\mathfrak{z}_{X_0}$. This proves that $V=\theta(\mathfrak{g}_{X_0})$.

Let ρ denote the representation $Z\rightarrow\operatorname{ad}Z$ $(Z\in\mathfrak{l})$ of $\mathfrak{l}$ on $\mathfrak{g}$. Since $\theta(\mathfrak{l})=\mathfrak{l}$, it follows from Lemma 27 that $\rho(\mathfrak{l})$ is a self-adjoint family of linear transformations on $\mathfrak{g}$. Therefore we can write $\mathfrak{g}=\sum_{1\leqq i\leqq r}\mathfrak{g}_i$ where $\mathfrak{g}_i\neq\{0\}$ is a subspace of $\mathfrak{g}$ which is invariant and irreducible under ρ and the sum is orthogonal. Furthermore we may assume that $\mathfrak{g}_1=\mathfrak{l}$.

Note that $\theta(X_0)=-Y_0$. Hence $U=\theta(\mathfrak{z}_{X_0})=\mathfrak{z}_{Y_0}$ where $\mathfrak{z}_{Y_0}$ is the centralizer of Y_0 in $\mathfrak{g}$. Similarly $V=\theta(\mathfrak{g}_{X_0})=[Y_0,\mathfrak{g}]$. Therefore $Y_0\in U$ and $H_0=-[Y_0,X_0]\in V$. Since $\mathfrak{z}_{X_0}$ is the kernel of the endomorphism $\rho(X_0)$ of $\mathfrak{g}$, it is clear that

$$\mathfrak{z}_{X_0}=\sum_{1\leqq i\leqq r}\mathfrak{z}_{X_0}\cap\mathfrak{g}_i.$$

Let $\lambda_i+1=\dim\mathfrak{g}_i$ so that $\lambda_i\geqq 0$ $(1\leqq i\leqq r)$. We know from Lemma 25 that $\dim(\mathfrak{z}_{X_0}\cap\mathfrak{g}_i)=1$. Let w'_i be a unit vector in $\mathfrak{z}_{X_0}\cap\mathfrak{g}_i$. Then $\rho(H_0)w'_i=\lambda_i w'_i$ from Lemma 25. Put $w_i=\theta(w'_i)$. Then

$$[H_0,w_i]=-\theta([H_0,w'_i])=-\lambda_i\theta(w'_i)=-\lambda_i w_i$$

and $w_1,\cdots,w_r$ is an orthonormal basis for $U=\theta(\mathfrak{z}_{X_0})$.

LEMMA 29. *Let $n=\dim\mathfrak{g}$. Then*

$$n=\sum_{1\leqq i\leqq r}(\lambda_i+1),\qquad r=\dim\mathfrak{z}_{X_0}.$$

This is obvious since $\mathfrak{g}=\sum_i\mathfrak{g}_i$.

Let $(u_1,\cdots,u_r)$ denote the Cartesian coordinates in U with respect to $(w_1,\cdots,w_r)$ so that $u=\sum_i u_iw_i$ $(u\in U)$. Then $\partial(w_i)=\partial/\partial u_i$.

LEMMA 30. *Let D denote the vector field on $\mathfrak{g}$ defined by $D_X = \partial(X)$ $(X \in \mathfrak{g})$. Then D is invariant and*

$$\Delta(D) = \sum_{1 \leq i \leq r} (1 + \tfrac{1}{2}\lambda_i) d_i$$

where $d_i = u_i \partial/\partial u_i$ $(1 \leq i \leq r)$.

Since $(D_X)^x = (\partial(X))^x = \partial(xX) = D_{xX}$ $(x \in G, X \in \mathfrak{g})$, it follows that D is invariant.

Now fix $u \in U'$. Then

$$\Gamma_{X_0+u}(H_0 \otimes 1) = [H_0, X_0 + u] = 2X_0 + [H_0, u].$$

Therefore

$$\Gamma_{X_0+u}(\tfrac{1}{2}H_0 \otimes 1 + 1 \otimes (u - \tfrac{1}{2}[H_0, u])) = X_0 + u.$$

Now $[H_0, U] \subset U$ since $U = \mathfrak{z}_{Y_0}$. Therefore

$$\Delta_u(D) = \partial(u) - \tfrac{1}{2}\partial([H_0, u]).$$

But $u = \sum_i u_i w_i$ and so

$$u - \tfrac{1}{2}[H_0, u] = \sum_i (1 + \tfrac{1}{2}\lambda_i) u_i w_i.$$

Therefore

$$\Delta_u(D) = \sum_i (1 + \tfrac{1}{2}\lambda_i) u_i \partial(w_i) \qquad (u \in U').$$

LEMMA 31. *Fix integers $m_i \geq 0$ $(1 \leq i \leq r)$, put*

$$d = \partial(w_1^{m_1} w_2^{m_2} \cdots w_r^{m_r})$$

and define D and d_i $(1 \leq i \leq r)$ as above. Then[9] *$\{d, d_i\} = m_i d$ and*

$$(d \circ \Delta(D)^*)_0 = -\sum_{1 \leq i \leq r} (1 + \tfrac{1}{2}\lambda_i)(m_i + 1) d.$$

It is clear that

$$\{\partial(w_j), d_i\} = \delta_{ji}\partial(w_i).$$

Hence for a fixed i,

$$\{d, d_i\} = \{\partial(w_i^{m_i}), d_i\} d'$$

where $d' = \prod_{j \neq i} \partial(w_j^{m_j})$. But it is obvious from the formula above that

$$\{\partial(w_i^{m_i}), d_i\} = m_i \partial(w_i^{m_i})$$

and therefore $\{d, d_i\} = m_i d$. Moreover $d_i^* = -(d_i + 1)$. Hence

$$\Delta(D)^* = -\sum_i (1 + \tfrac{1}{2}\lambda_i)(d_i + 1)$$

from Lemma 30. Also

---

[9] If D_1, D_2 are two differential operators, we denote $D_1 \circ D_2 - D_2 \circ D_1$ by $\{D_1, D_2\}$.

$$dd_i = \{d, d_i\} + d_i d = (d_i + m_i) d.$$

Therefore since $(d_i d)_0 = 0$, we have $(dd_i)_0 = m_i d$. Hence

$$(d \circ \Delta(D)^*)_0 = -\sum_i (1 + \tfrac{1}{2}\lambda_i)(m_i + 1) d.$$

Let $\mathfrak{z}_{\mathfrak{l}}$ be the centralizer of $\mathfrak{l}$ in $\mathfrak{g}$.

LEMMA 32. *The following four conditions are mutually equivalent.*

(1) ω *is identically zero on* $U = \mathfrak{z}_{Y_0}$;

(2) ω *is identically zero on* $\theta(U) = \mathfrak{z}_{X_0}$;

(3) $\lambda_i > 0$ $(1 \leqq i \leqq r)$;

(4) $\mathfrak{z}_{\mathfrak{l}} = \{0\}$.

We know that ω is invariant under θ. Hence (1) and (2) are equivalent. Since X_0, Y_0 generate $\mathfrak{l}$ as a Lie algebra, it is obvious that $\mathfrak{z}_{\mathfrak{l}} = \mathfrak{z}_{X_0} \cap \mathfrak{z}_{Y_0}$. Let ρ_i denote the representation of $\mathfrak{l}$ on $\mathfrak{g}_i$ defined under ρ $(1 \leqq i \leqq r)$. Then ρ_i is the zero representation if and only if $\lambda_i = 0$. On the other hand $\dim \mathfrak{z}_{\mathfrak{l}}$ is the number of times the zero representation of $\mathfrak{l}$ occurs in the complete reduction of ρ. Hence

$$\dim \mathfrak{z}_{\mathfrak{l}} = \sum_{\lambda_i = 0} (\lambda_i + 1).$$

This proves that (3) and (4) are equivalent.

Now assume (2) holds. Then $B(X, Y) = 0$ for $X, Y \in \mathfrak{z}_{X_0}$ and therefore $\mathfrak{z}_{X_0}$ and $\theta(\mathfrak{z}_{X_0}) = \mathfrak{z}_{Y_0}$ are mutually orthogonal. But then $\mathfrak{z}_{\mathfrak{l}} = \mathfrak{z}_{X_0} \cap \mathfrak{z}_{Y_0} = \{0\}$. Therefore (2) implies (4).

Conversely suppose (3) holds. Then $\mathfrak{z}_{Y_0} = U = \sum_i \boldsymbol{R} w_i$ and

$$B(w_i, w_j) = -\langle \theta(w_i), w_j \rangle \qquad (1 \leqq i, j \leqq r).$$

We have seen that $w'_i = \theta(w_i)$ and w_j are eigenvectors of $\rho(H_0)$ belonging to the eigenvalues λ_i and $-\lambda_j$ respectively. It follows from (3) that these two eigenvalues are distinct. Since $\rho(H_0)$ is self-adjoint (by Lemma 27), the vectors w'_i and w_j are orthogonal. Hence $B(w_i, w_j) = 0$ $(1 \leqq i, j \leqq r)$. This shows that (3) implies (1) and so the lemma is proved.

Now consider the differential operator $\Delta(\partial(\omega))$ on U'. It is clear (see the remark at the end of § 7) that its degree is $\leqq 2$. As usual $\Delta_0(\partial(\omega))$ denotes its local expression at zero.

LEMMA 33. *The homogeneous part of degree 2 of* $\Delta_0(\partial(\omega))$ *is zero if and only if* $\mathfrak{z}_{\mathfrak{l}} = \{0\}$.

Choose a base v_i $(1 \leqq i \leqq s)$ for V. Then $[X_0, v_i]$, w_j $(1 \leqq i \leqq s, 1 \leqq j \leqq r)$ is a base for $\mathfrak{g}$. Hence

$$\omega = \sum_{1 \leqq i, j \leqq s} a^{ij}[X_0, v_i][X_0, v_j] + \sum_{1 \leqq i, j \leqq r} b^{ij} w_i w_j$$
$$+ \sum_{1 \leqq i \leqq s} \sum_{1 \leqq j \leqq r} c^{ij}[X_0, v_i] w_j$$

where $a^{ij} = a^{ji}$, $b^{ij} = b^{ji}$ and c^{ij} are in $\boldsymbol{R}$. Let $\cdot$ denote product in $\mathfrak{S}$. Then

$$\Gamma_{X_0}(v_i \cdot v_j \otimes 1) = [X_0, v_i][X_0, v_j] - [v_i, [X_0, v_j]],$$

and

$$\Gamma_{X_0}(v_i \otimes w_j) = -[X_0, v_i] w_j + [v_i, w_j].$$

Therefore if λ denotes, as usual, the canonical mapping of $S(\mathfrak{g}_c)$ onto $\mathfrak{S}$, we get

$$\sum_{1 \leqq i, j \leqq s} a^{ij} \Gamma_{X_0}(\lambda(v_i v_j) \otimes 1) - \sum_{1 \leqq i \leqq s} \sum_{1 \leqq j \leqq r} c^{ij} \Gamma_{X_0}(v_i \otimes w_j)$$
$$= \sum_{i,j} a^{ij}[X_0, v_i][X_0, v_j] - \sum_{i,j} a^{ij}[v_i, [X_0, v_j]]$$
$$+ \sum_{i,j} c^{ij}[X_0, v_i] w_j - \sum_{i,j} c^{ij}[v_i, w_j]$$
$$= \omega - \sum_{1 \leqq i, j \leqq r} b^{ij} w_i w_j + \mu$$

where μ is an element of $S(\mathfrak{g}_c)$ of degree $\leqq 1$. This shows that the homogeneous part of degree 2 of $\Delta_0(\partial(\omega))$ is $\sum_{i,j} b^{ij} \partial(w_i w_j) = \sum_{i,j} b^{ij} \partial^2 / \partial u_i \partial u_j$.

So it remains to show that $b^{ij} = 0$ $(1 \leqq i, j \leqq r)$ if and only if $\mathfrak{z}_I = \{0\}$. Now $\mathfrak{g}_{X_0}$ is orthogonal to $U = \theta(\mathfrak{z}_{X_0})$. Therefore $B(X, Z) = 0$ for $X \in \mathfrak{g}_{X_0}$ and $Z \in \mathfrak{z}_{X_0}$. Hence if $v \in V$, $[X_0, v]$, regarded as a linear function on $\mathfrak{g}$, vanishes identically on $\mathfrak{z}_{X_0}$. Therefore if $X \in \mathfrak{z}_{X_0}$,

$$\omega(X) = \sum_{1 \leqq i, j \leqq r} b^{ij} B(w_i, X) B(w_j, X).$$

Now $w_1, \cdots, w_r$ is an orthonormal base for U. Hence $\theta(w_i)$ $(1 \leqq i \leqq r)$ is a base for $\theta(U) = \mathfrak{z}_{X_0}$. Let $X = \sum_{1 \leqq i \leqq r} t_i \theta(w_i)$ $(t \in \boldsymbol{R})$. Then it follows from the above orthonormality that $B(w_i, X) = -t_i$. Hence

$$\omega(X) = \sum_{1 \leqq i, j \leqq r} b^{ij} t_i t_j.$$

The required result is now obvious from Lemma 32.

COROLLARY 1. *Suppose* $\mathfrak{z}_I \neq \{0\}$. *Then if* $z \in S(U_c)$ *and*

$$(\partial(z) \circ \Delta(\partial(\omega))^*)_0 = 0,$$

we can conclude that $z = 0$.

For otherwise suppose $z \neq 0$. Let m be the degree of z and z_0 the homogeneous part of degree m of z. Put $\Delta = \Delta(\partial(\omega))$. Then

$$\Delta = \sum_{1 \leqq i,j \leqq r} a_{ij}\partial^2/\partial u_i \partial u_j + \sum_{1 \leqq i \leqq r} a_i \partial/\partial u_i + a$$

where $a_{ij} = a_{ji}$, a_i and a are C^∞ functions on U'. It follows from Lemma 33 that

$$\sum_{i,j} a_{ij}(0)\partial^2/\partial u_i \partial u_j \neq 0.$$

Moreover it is clear that

$$\Delta^* = \sum_{i,j} a_{ij}\partial^2/\partial u_i \partial u_j + \sum_i b_i \partial/\partial u_i + b$$

where b_i and b are also C^∞ functions on U'. Since

$$\partial(z) \circ \Delta^* = \Delta^* \circ \partial(z) + \{\partial(z), \Delta^*\},$$

it is clear that the homogeneous part of degree $m+2$ of $(\partial(z) \circ \Delta^*)_0$ is

$$\sum_{i,j} a_{ij}(0)\partial^2/\partial u_i \partial u_j \circ \partial(z_0) \neq 0.$$

Therefore $(\partial(z) \circ \Delta^*)_0 \neq 0$ and so we get a contradiction.

COROLLARY 2. *If* $\mathfrak{z}_{\mathfrak{l}} \neq \{0\}$, *then* $z = 0$ *in the corollary of Lemma* 23 *and therefore* $\sigma_T = 0$.

For since $\partial(\omega)T = 0$, we conclude from Theorem 3 that $\Delta\sigma_T = 0$ where $\Delta = \Delta(\partial(\omega))$. But then it follows from the corollary of Lemma 23 that $\beta(0; \partial(z) \circ \Delta^*) = 0$ for every $\beta \in C_c^\infty(U_0)$. This however implies that $(\partial(z) \circ \Delta^*)_0 = 0$. Our assertion now follows from Corollary 1 above.

Thus in order to complete the proof of Lemma 17, we may, from now on, assume that $\mathfrak{z}_{\mathfrak{l}} = \{0\}$. Then, by Lemma 32, ω vanishes identically on $U = \theta(\mathfrak{z}_{X_0})$ and $\mathfrak{z}_{X_0}$. Therefore

$$\omega(X_0 + u) = \omega(X_0) + \omega(u) + 2B(X_0, u) = 2B(X_0, u)$$

for $u \in U$. Now w_1 could be chosen to be any unit vector such that $\theta(w_1) \in \mathfrak{g}_1 \cap \mathfrak{z}_{X_0} = \boldsymbol{R}X_0$. (We recall that $\mathfrak{g}_1 = \mathfrak{l}$.) So we may take

$$w_1 = \| Y_0 \|^{-1} Y_0.$$

Then

$$u_1 = \langle w_1, u \rangle = -B(\theta(w_1), u) = \| Y_0 \|^{-1} B(X_0, u).$$

Hence $\omega(X_0 + u) = 2cu_1$ $(u \in U)$ where $c = \| X_0 \| = \| Y_0 \|$. We also note that $\dim \mathfrak{g}_1 = \dim \mathfrak{l} = 3$ and therefore $\lambda_1 = 2$.

Now ω vanishes identically on $\mathfrak{N}$. Therefore, by Lemma 21, there exists an integer $k \geqq 0$ such that $\omega^k T = 0$ around X_0. On the other hand Lemma 19 would be proved if we can show that $T = 0$ around X_0. Therefore it would clearly be sufficient to obtain the following result.

LEMMA 34. *Fix a point $X_0 \in \mathfrak{N}$ and an integer $k \geqq 0$ and let q be the topological dimension of the orbit $X_0{}^G$. Let T be a locally invariant distribution on some open neighborhood Ω of X_0 in $\mathfrak{g}$ such that*

1) $\omega^k T = 0$, $\partial(\omega) T = 0$
2) $\operatorname{Supp} T \subset \mathfrak{N}_q \cap \Omega$.

Then $T = 0$ around X_0.

The case $k = 0$ being trivial, we assume $k \geqq 1$ and use induction. Also, in view of our results above, we may suppose that $q \geqq 1$ and $\mathfrak{z}_{\mathfrak{l}} = \{0\}$. Finally, we may assume that $X_0 \in \operatorname{Supp}(\omega^{k-1} T)$ for otherwise the assertion of the lemma would follow immediately from the induction hypothesis.

15. Some algebraic lemmas. We now study some simple algebraic properties of the differential operator D of Lemma 30.

LEMMA 35. *Let $p \in S_m(\mathfrak{g}_c)$. Then $D \circ p = p(D + m)$ and $\partial(p) \circ D = (D + m) \circ \partial(p)$.*

For any $f \in C^\infty(\mathfrak{g})$ and $t \in \boldsymbol{R}$, let f_t denote the function $X \to f(tX)$ $(X \in \mathfrak{g})$. Then

$$f(X; D) = f(X; \partial(X)) = \{df(tX)/dt\}_{t=1} \qquad (X \in \mathfrak{g})$$

and therefore $Df = (df_t/dt)_{t=1}$. Hence

$$D(pf) = \{d(pf)_t/dt\}_{t=1}.$$

But it is obvious that $p_t = t^m p$ and therefore $(pf)_t = p_t f_t = t^m p f_t$. Therefore

$$D(pf) = \{d(t^m p f_t)/dt\}_{t=1} = mpf + pDf$$

and this proves the first statement.

Now we come to the second part. Let $X_1, \cdots, X_m \in \mathfrak{g}$. Then

$$\begin{aligned} f_t(X; \partial(X_1 X_2 \cdots X_m)) &= \{\partial^m f_t(X + t_1 X_1 + \cdots + t_m X_m)/\partial t_1 \cdots \partial t_m\}_{t_1 = \cdots = t_m = 0} \\ &= \{t^m \partial^m f(tX + t_1 X_1 + \cdots + t_m X_m)/\partial t_1 \cdots \partial t_m\}_{t_1 = \cdots = t_m = 0} \\ &= t^m f(tX; \partial(X_1 \cdots X_m)). \end{aligned}$$

This shows that $\partial(p) f_t = t^m (\partial(p) f)_t$ for $p \in S_m(\mathfrak{g}_c)$. Therefore

$$\partial(p)(Df) = (d(\partial(p)f_t)/dt)_{t=1} = m\partial(p)f + (d(\partial(p)f)_t/dt)_{t=1}$$
$$= m\partial(p)f + D(\partial(p)f).$$

This completes the proof of Lemma 35.

LEMMA 36. *Put* $n = \dim \mathfrak{g}$. *Then*

$$\{\partial(\omega), \omega\} = 4(D + n/2).$$

Let X_i $(1 \leqq i \leqq n)$ be a base for $\mathfrak{g}$ and X^i $(1 \leqq i \leqq n)$ the dual base so that $B(X^i, X_j) = \delta_j^i$. Let $X^i = \sum_j g^{ij} X_j$ and $X_i = \sum_j g_{ij} X^j$ $(g_{ij}, g^{ij} \in \boldsymbol{R})$. Then $g^{ij} = B(X^i, X^j)$, $g_{ij} = B(X_i, X_j)$ and the matrices (g^{ij}) and (g_{ij}) are inverse to each other. Moreover

$$\partial(X_i)X_j = B(X_i, X_j) = g_{ij}.$$

Now let $X = \sum_i t^i X_i$ $(t^i \in \boldsymbol{R})$. Then

$$\omega(X) = \sum_{i,j} t^i t^j B(X_i, X_j) = \sum_{i,j} g_{ij} t^i t^j.$$

But $t^i = B(X^i, X)$. Hence

$$\omega = \sum_{i,j} g_{ij} X^i X^j = \sum_{i,j} g^{ij} X_i X_j$$

and therefore

$$\{\partial(X_k), \omega\} = \sum_{i,j} g^{ij}(g_{ki} X_j + X_i g_{kj})$$
$$= 2X_k.$$

Hence

$$\{\partial(\omega), \omega\} = \sum_k (\{\partial(X^k), \omega\} \circ \partial(X_k) + \partial(X^k) \circ \{\partial(X_k), \omega\})$$
$$= 2 \sum_k (X^k \partial(X_k) + \partial(X_k) \circ X^k)$$
$$= 2n + 4 \sum_k X^k \partial(X_k).$$

So it remains to verify that $\sum_k X^k \partial(X_k) = D$. This is seen as follows. If $X \in \mathfrak{g}$, then

$$(\sum_k X^k \partial(X_k))_X = \sum_k B(X^k, X)\partial(X_k) = \partial(X) = D_X$$

since $B(X^k, X) = t^k$ if $X = \sum_k t^k X_k$ $(t^k \in \boldsymbol{R})$.

COROLLARY. *Let* l *be any integer* $\geqq 1$. *Then*

$$\{\partial(\omega), \omega^l\} = 4l\omega^{l-1}(D + n/2 + l - 1).$$

If $l = 1$, this follows from Lemma 36. So assume $l \geqq 2$ and use induction. Then $\omega^l = \omega\omega^{l-1}$ and therefore

$$\begin{aligned}\{\partial(\omega),\omega^l\} &= \{\partial(\omega),\omega\}\circ\omega^{l-1}+\omega\{\partial(\omega),\omega^{l-1}\}\\ &= 4(D+n/2)\circ\omega^{l-1}+4(l-1)\omega^{l-1}(D+n/2+l-2)\end{aligned}$$

by induction hypothesis. But $(D+n/2)\circ\omega^{l-1}=\omega^{l-1}(D+n/2+2l-2)$ from Lemma 35. From this the required result follows immediately.

16. Completion of the proof of Theorem 4. Now we return to the proof of Lemma 34. Put $T_0=(D+n/2+k-1)T$. Since D is invariant, T_0 is a locally invariant distribution on Ω and $\operatorname{Supp} T_0 \subset \operatorname{Supp} T$. Moreover

$$\begin{aligned}0 &= \partial(\omega)(\omega^kT)-\omega^k(\partial(\omega)T)\\ &= 4k\omega^{k-1}T_0\end{aligned}$$

from the corollary of Lemma 36. This shows that $\omega^{k-1}T_0=0$. Moreover

$$\begin{aligned}\partial(\omega)T_0 &= \partial(\omega)(D+n/2+k-1)T\\ &= (D+n/2+k+1)\partial(\omega)T=0\end{aligned}$$

from Lemma 35. Thus T_0 satisfies all the conditions of Lemma 34 with k replaced by $k-1$. Hence by induction hypothesis $T_0=0$ around X_0. So by replacing Ω, if necessary, by a smaller open neighborhood of X_0, we may assume that $T_0=0$ on Ω. Therefore $(D+n/2+k-1)T=0$.

Let δ_0 denote the distribution $\beta\to\beta(0)$ $(\beta\in C_c^\infty(U_0))$ on U_0. Then, from Lemmas 20 and 23, we can choose $z\in S(U_c)$ such that $\sigma_T=\partial(z)\delta_0$. If $p\in I(\mathfrak{g}_0)$, it is obvious that $\Delta(p)=q$ where q is the polynomial function $u\to p(X_0+u)$ $(u\in U')$. Since $\omega(X_0+u)=2cu_1$ $(u\in U)$, it follows from Theorem 3 that $\sigma_{\omega^m T}=(2cu_1)^m\partial(z)\delta_0$ for any integer $m\geqq 0$. Therefore since $\omega^kT=0$, $X_0\in\operatorname{Supp}(\omega^{k-1}T)$ and $(D+n/2+k-1)T=0$, we obtain the following result by taking into account Theorems 2 and 3.

LEMMA 37. *$u_1^k\partial(z)\delta_0=0$ but $u_1^{k-1}\partial(z)\delta_0\neq 0$. Moreover*

$$(\Delta(D)+n/2+k-1)\partial(z)\delta_0=0.$$

The following lemma is proved by a simple calculation.

LEMMA 38. *$u_1^p\partial(w_1^{m_1}w_2^{m_2}\cdots w_r^{m_r})\delta_0=0$ if $p>m_1$ and*

$$=(-1)^p\frac{m_1!}{(m_1-p)!}\,\partial(w_1^{m_1-p}w_2^{m_2}\cdots w_r^{m_r})\delta_0$$

if $p\leqq m_1$.

Put $d(m)=\partial(w_1^{m_1}\cdots w_r^{m_r})$ for $m=(m_1,\cdots,m_r)$ and suppose

$$\partial(z)=\sum_m c(m)d(m)$$

where $c(m)$ are complex numbers which are zero for all m except a finite number. Then

$$0 = u_1^k\partial(z)\delta_0$$
$$= \sum_{m_1 \geqq k} \sum_{m_2, \cdots, m_r \geqq 0} c(m)(-1)^k(m_1!/(m_1-k)!)\partial(w_1^{m_1-k}w_2^{m_2}\cdots w_r^{m_r})\delta_0$$

and therefore it follows from Lemma 20 that $c(m) = 0$ if $m_1 \geqq k$. Similarly

$$u_1^{k-1}\partial(z)\delta_0$$
$$= \sum_{m_1 \geqq k-1} \sum_{m_2, \cdots, m_r \geqq 0} c(m)(-1)^{k-1}(m_1!/(m_1-k+1)!)\partial(w_1^{m_1-k+1}w_2^{m_2}\cdots w_r^{m_r})\delta_0.$$

Therefore it is obvious from the above result and Lemma 37 that $c(m) \neq 0$ for some m with $m_1 = k-1$.

On the other hand

$$\sum_m c(m)(\Delta(D) + n/2 + k - 1)d(m)\delta_0 = 0$$

from Lemma 37. It follows from Lemma 31 that

$$\Delta(D)d(m)\delta_0 = -\sum_{1 \leqq i \leqq r}(1 + \tfrac{1}{2}\lambda_i)(m_i + 1)d(m)\delta_0.$$

Therefore

$$\sum_m c(m)\{n/2 + k - 1 - \sum_i (1 + \tfrac{1}{2}\lambda_i)(m_i + 1)\}d(m)\delta_0 = 0$$

and this implies that $c(m) = 0$ unless

$$\sum_i (1 + \tfrac{1}{2}\lambda_i)(m_i + 1) = n/2 + k - 1.$$

But $\sum_i(1 + \frac{1}{2}\lambda_i) = \frac{1}{2}(n + r)$ from Lemma 29. Therefore if $c(m) \neq 0$, we must have

$$\sum_i (1 + \tfrac{1}{2}\lambda_i)m_i + r/2 = k - 1.$$

Since $r \geqq 1$ and $\lambda_1 = 2$, this implies that $2m_1 < k - 1$. However we have seen above that $c(m) \neq 0$ for some m with $m_1 = k - 1$. This contradiction completes the proof of Lemma 34 and therefore also of Theorem 4.

17. Proof of Theorem 5. It is now easy to generalize Theorem 4 as follows.

THEOREM 5. *Let T be a distribution on an open subset Ω of $\mathfrak{g}$. Assume that*

1) *T is locally invariant,*
2) $\operatorname{Supp} T \subset \mathfrak{N} \cap \Omega$,

3) *T satisfies an equation of the form* $\sum_{0 \leqq k \leqq r} c_k \partial(\omega^k) T = 0$ *where* c_k *are complex numbers not all zero.*

Then $T = 0$.

Let $\mathfrak{X}$ be the space of distributions on Ω spanned over $\boldsymbol{C}$ by $\partial(\omega^k)T$ $(k \geqq 0)$. Then, in view of condition (3), $\dim \mathfrak{X} < \infty$. Since $T \in \mathfrak{X}$, it would be enough to prove that $\mathfrak{X} = \{0\}$. So let us assume that $\mathfrak{X} \neq \{0\}$. Since $\partial(\omega)$ defines a linear transformation in $\mathfrak{X}$, we can choose $T_0 \neq 0$ in $\mathfrak{X}$ such that $\partial(\omega)T_0 = cT_0$ for some $c \in \boldsymbol{C}$. It is clear that every distribution in $\mathfrak{X}$ satisfies conditions (1) and (2). Fix a point X_0 in the support of T_0. Since ω vanishes identically on $\mathfrak{N}$, it follows from Lemma 21 that we can choose an open neighborhood Ω_0 of X_0 in Ω and an integer $k \geqq 0$ such that $\omega^k T_0 = 0$ on Ω_0.

LEMMA 39. *For any differential operator* D *on* Ω_0, *define*

$$d_\omega D = \{\partial(\omega), D\}.$$

Then if $DT_0 = 0$ *on* Ω_0, *the same holds for* $(d_\omega D)T_0$.

Let $D' = d_\omega D$. Then

$$\begin{aligned} D'T_0 &= \partial(\omega)(DT_0) - D(\partial(\omega)T_0) \\ &= (\partial(\omega) - c)(DT_0) = 0 \end{aligned}$$

on Ω_0, since $\partial(\omega)T_0 = cT_0$.

Therefore we conclude by induction on r that $(d_\omega{}^r \omega^k)T_0 = 0$ on Ω_0. But $d_\omega{}^{2k}\omega^k = 2k!2^{2k}\partial(\omega^k)$ (see [4(c), p. 99]). Hence $\partial(\omega^k)T = 0$ on Ω_0. However $\partial(\omega^k)T_0 = c^k T_0$. Therefore since $X_0 \in \operatorname{Supp} T_0$, we must have $c = 0$. But then $T_0 = 0$ from Theorem 4 and so we get a contradiction.

18. Appendix. The results of this section were pointed out to me by A. Borel and they are reproduced here with his permission.

We use the terminology of [2, §§1-2].

LEMMA 40. *Let* $G \subset \boldsymbol{GL}(n, \boldsymbol{R})$ *be a real algebraic Lie group.* G' *an open subgroup of* G, $f\colon G \to \boldsymbol{GL}(m, \boldsymbol{R})$ *a rational representation of* G *and* $X = f(G') \cdot x$ $(x \in \boldsymbol{R}^m)$ *an orbit of* G'. *Let* q *be the topological dimension of* X. *Then* X *is open in*[10] $\mathrm{Cl}\, X$ *and* $\mathrm{Cl}\, X \cap {}^cX$ *is the union of orbits of* G' *of dimension* $< q$.

---

[10] For any set U, $\mathrm{Cl}\, U$ stands for the topological closure of U and ${}^cU$ for the complement of U.

Let $G_c \subset \mathbf{GL}(n, \mathbf{C})$ be the smallest complex algebraic group containing G. Then f extends uniquely to a rational representation $f_c\colon G_c \rightarrow \mathbf{GL}(m, \mathbf{C})$ [3(b), Prop. 4, p. 109]. Consider the orbit $Y = f_c(G_c) \cdot x$ in $\mathbf{C}^m$. On each irreducible component of G_c, the map $g \rightarrow f_c(g) \cdot x$ is rational and regular and therefore Y is a complex manifold of complex dimension q. Since the image of an irreducible variety under a rational regular map contains a non-empty Zariski-open subset of its Zariski closure, it follows that Y contains an open everywhere dense subset of $\mathrm{Cl}\, Y$ and that $V = \mathrm{Cl}\, Y$ is an algebraic set of algebraic dimension q. By homogeneity, Y is open in V and consists of simple points and $V \cap {}^cY$ is an algebraic set of algebraic dimension $q' < q$. Also since G is Zariski-dense in G_c, by definition of G_c, it is clear that V is the smallest algebraic set containing $f(G) \cdot x$.

Let $a \in Y_{\mathbf{R}}$. Then a is simple on V and there exists a neighborhood U of a in $\mathbf{R}^m$ such that $U \cap V_{\mathbf{R}}$ is a connected manifold of dimension q (see Whitney [10, § 11]). On the other hand the isotropy group of a in G' is contained in the set of real points of an algebraic group of complex dimension $r - q$ where $r = \dim G$. Hence its dimension is $\leqq r - q$. Therefore $f(G') \cdot a$ has dimension $\geqq q$ and so it contains $U \cap V_{\mathbf{R}}$ if U is taken sufficiently small. This shows that a G'-orbit in $Y_{\mathbf{R}}$ is open in $V_{\mathbf{R}}$. Therefore ${}^cX \cap Y_{\mathbf{R}}$, being a union of such orbits, is open in $V_{\mathbf{R}}$. Hence X is closed in $Y_{\mathbf{R}}$. This proves that $\mathrm{Cl}\, V \cap Y_{\mathbf{R}} = X$ and therefore $\mathrm{Cl}\, X \subset X \cup (V_{\mathbf{R}} \cap {}^cY_{\mathbf{R}})$. Since $V \cap {}^cY$ is an algebraic set of algebraic dimension $q' < q$, the topological dimension of $V_{\mathbf{R}} \cap {}^cY_{\mathbf{R}} = (V \cap {}^cY)_{\mathbf{R}}$ is $\leqq q'$. (This follows by induction on dimension from [10, Theorem 1 and Lemma 9].) Since $(V \cap {}^cY)_{\mathbf{R}}$ is stable under G, our assertion follows immediately.

COROLLARY. *Let A be a real algebraic subset of $\mathbf{R}^m$ which is the union of a finite number of G'-orbits. Let A_q denote the union of all G'-orbits in A of topological dimension $\leqq q$. Then A_q is closed and each G'-orbit of dimension q is open in A_q.*

It is clear from the above lemma that A_q is closed. Let $X = f(G') \cdot x$ $(x \in A)$ be an orbit of dimension q. Then by Lemma 40, $A_q \cap {}^cX$ is closed and hence X is open in A_q.

It is now easy to give a proof of Lemma 17. Take G to be the group of all automorphisms of $\mathfrak{g}$ and G' the connected component of 1 in G. Then G' is an open subgroup of the real algebraic Lie group G. Let $n = \dim \mathfrak{g}$. Define polynomial functions p_i $(0 \leqq i < n)$ on $\mathfrak{g}$ by the relation

$$\det(t - \operatorname{ad} X) = \sum_{0 \leqq i < n} p_i(X) t^i + t^n \qquad (X \in \mathfrak{g}),$$

where t is an indeterminate. Then obviously $\mathfrak{N}$ is the set of all common zeros of p_i $(0 \leqq i < n)$ in $\mathfrak{g}$. Hence $\mathfrak{N}$ is a real algebraic subset of $\mathfrak{g}$. Lemma 17 is an immediate consequence of the above corollary if we take $A = \mathfrak{N}$.

INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] F. A. Berezin, "Laplace operators on semisimple Lie groups," *American Mathematical Society Translations*, vol. 21 (1962), pp. 239-339.

[2] A. Borel and Harish-Chandra, "Arithmetic subgroups of algebraic groups," *Annals of Mathematics*, vol. 75 (1962), pp. 485-535.

[3] C. Chevalley, (a) *Theory of Lie groups*, Princeton University Press, 1946.

(b) *Groupes Algébriques*, Actualités Scientifiques et Industrielles 1152, Paris, 1951.

[4] Harish-Chandra, (a) "Representations of a semisimple Lie group on a Banach space," *Transactions of the American Mathematical Society*, vol. 75 (1953), pp. 185-243.

(b) "Representations of semisimple Lie groups, III," *Transactions of the American Mathematical Society*, vol. 76 (1954), pp. 234-253.

(c) "Differential operators on a semisimple Lie algebra," *American Journal of Mathematics*, vol. 79 (1957), pp. 87-120.

(d) "The characters of semisimple Lie groups," *Transactions of the American Mathematical Society*, vol. 83 (1956), pp. 98-163.

(e) "Invariant eigendistributions on semisimple Lie groups," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 117-123.

[5] N. Jacobson, "Completely reducible Lie algebras of linear transformations," *Proceedings of the American Mathematical Society*, vol. 2 (1951), pp. 105-113.

[6] B. Kostant "The principal three-dimensional subgroup and the Betti numbers of a complex simple Lie group," *American Journal of Mathematics*, vol. 81 1959), pp. 973-1032.

[7] D. Montgomery and L. Zippin, *Topological transformation groups*, Interscience Tracts in Pure and Applied Mathematics, I, New York, 1955.

[8] G. D. Mostow, *Some new decomposition theorems for semisimple groups*, Memoirs of the American Mathematical Society, vol. 14 (1955), pp. 31-54.

[9] L. Schwartz, *Théorie des distributions* I, Paris, Hermann, 1950.

[10] H. Whitney, "Elementary structure of real algebraic varieties," *Annals of Mathematics*, vol. 66 (1957), pp. 545-556.

COMMUTATORS, ABSOLUTELY CONTINOUS SPECTRA, AND SINGULAR INTEGRAL OPERATORS.*

By C. R. PUTNAM.

1. Introduction. Let $\mathfrak{H}$ denote a Hilbert space of elements $x, y, \cdots$, with the inner product (x, y). Only bounded operators $A, B, \cdots$, on $\mathfrak{H}$ will be considered. For such an operator A, let $\| A \|$ denote the usual norm defined by $\| A \| = \sup \| Ax \|$, with $\| x \| = 1$, and let $\operatorname{sp}(A)$ denote the spectrum of A. If A is self-adjoint with the spectral resolution

$$A = \int \lambda dE(\lambda), \tag{1}$$

then the set, $\mathfrak{H}_a$, of elements x in $\mathfrak{H}$ for which $\| E(\lambda)x \|^2$ is an absolutely continuous function of λ, is a subspace of $\mathfrak{H}$ which reduces the operator A; cf. Kato [4], p. 240, Kuroda [7], p. 436, Halmos [17], p. 104. The operator A will be called absolutely continuous if $\mathfrak{H} = \mathfrak{H}_a$.

For any pair of operators A and B with the commutator

$$AB - BA = C, \tag{2}$$

let $\mathfrak{M}_B$ be the smallest subspace of $\mathfrak{H}$ which reduces $\operatorname{Im}(B) = (B - B^*)/2i$ and which contains $\mathfrak{R}(C)$, the range of C. Let d_B denote the difference between the maximum and minimum points of $\operatorname{sp}(\operatorname{Im}(B))$. The symbol "meas" will refer to ordinary one-dimensional Lebesgue measure. There will be proved the following

THEOREM 1. *Let A be self-adjoint, B be arbitrary, and the commutator C of (2) be self-adjoint and satisfy*

$$C \geqq 0 \quad (or\ C \leqq 0). \tag{3}$$

(I) *Then*

$$\pi \| C \| \leqq (d_B/2) \operatorname{meas} \operatorname{sp}(A), \tag{4}$$

and hence

$$\pi \| C \| \leqq \| B \| \operatorname{meas} \operatorname{sp}(A). \tag{4'}$$

Received June 28, 1963.

* This work was supported by the National Science Foundation research grant NSF-G18915.

(II)* *In addition,*

$$\mathfrak{N} \subset \mathfrak{H}_a, \tag{5}$$

where $\mathfrak{N}$ denotes the smallest subspace reducing A and containing $\mathfrak{M}_B$.

There follows the

COROLLARY *of* (I). *If A is self-adjoint, B is arbitrary, and C of* (2) *is self-adjoint and satisfies* (3), *then*

$$\| C \| \leqq (2/\pi) \| A \| \| B \|. \tag{6}$$

The inequalities (4), (4′) and (6) are optimal in the sense that there exist examples (see Section 5) fulfilling the hypothesis of Theorem 1 for which equality holds in (6), hence also in (4) and (4′). On the other hand, it is clear that for arbitrary A and B the inequality $\| AB - BA \| \leqq 2 \| A \| \| B \|$ is the best possible. In fact, as examples with finite matrices shows, it can happen that A and iB (hence also $AB - BA$) are self-adjoint and that $0 < \| AB - BA \| = 2 \| A \| \| B \|$ holds. Thus the restriction (3) is essential for the general validity of (6). In this connection, see the discussion of [12], pp. 215, 217; it can be noted that the above Corollary answers the question raised on p. 217.

A theorem similar to (I) but where the role of the self-adjoint A is played by a unitary operator U was given in [12], p. 217. Since this result will be needed in the proof of (I) it will be stated as a

LEMMA. *Let J denote a bounded self-adjoint operator, let U be unitary, and suppose that $UJU^* - J = D \geqq 0$ (or $\leqq 0$). Then*

$$\operatorname{meas} \operatorname{sp}(U) \geqq 2\pi \| D \| / d, \tag{7}$$

where d denotes the distance between the maximum and minimum points of $\operatorname{sp}(J)$.

The proofs of (I) and (II) will be given in Sections 2 and 3. Applications to the investigation of the spectra of the singular self-adjoint integral operators A on $\mathfrak{H} = L^2(a, b)$ of the form

$$Ax = h(t)x(t) + (1/i\pi)\phi(t) \int_a^b \phi^*(s)(s - t)^{-1} x(s)\, ds \tag{8}$$

will be given in Sections 4 and 5.

Concerning (8) and similar operators, see [1], [2], [3], [5], [6], [8], [9], [13], [14], [15], [16] and their bibliographies.

---

* See Appendix.

The author is indebted to L. de Branges for valuable discussions concerning portions of the subject of the present paper and for making available the paper [9] of J. D. Pincus.

2. Proof of (I). On taking adjoints in (2) it is seen that

$$AJ - JA = -iC, \text{ where } J = (B - B^*)/2i. \tag{9}$$

Hence $(A + iI)J - J(A + iI) = -iC$ and, on applying $(A + iI)^{-1}$ on the right and left in (9), one obtains

$$J(A + iI)^{-1} - (A + iI)^{-1}J = -i(A + iI)^{-1}C(A + iI)^{-1}. \tag{10}$$

If U denotes the Cayley transform of A,

$$U = (A - iI)(A + iI)^{-1} \qquad (= I - 2i(A + iI)^{-1}), \tag{11}$$

it follows from (10) that $JU - UJ = -2(A + iI)^{-1}C(A + iI)^{-1}$, and consequently,

$$UJU^* - J = 2Q^*CQ, \text{ where } Q = (A - iI)^{-1}. \tag{12}$$

If now the present J defined in (9) is identified with that occurring in the Lemma and if $D = 2Q^*CQ$, it is seen that $D \geqq 0$ (or $D \leqq 0$) and hence, by (7),

$$\text{meas sp}(U) \geqq 2\pi \| D \|/d, \text{ with } d = d_B. \tag{13}$$

But $\| Q^*CQ \| = \sup(Q^*CQx, x) \geqq \| C \| \inf \| Qx \|^2$, where $\| x \| = 1$; also, by (1) and the definition of Q in (12), $Q = \int (\lambda - i)^{-1} dE(\lambda)$, and so $\| Qx \|^2 \geqq (1 + \| A \|^2)^{-1}$. Consequently, relation (13) implies

$$\text{meas sp}(U) \geqq 4\pi \| C \|/d(1 + \| A \|^2). \tag{14}$$

If α is defined by

$$e^{i\alpha} = (\lambda - i)(\lambda + i)^{-1}, \text{ where } -\infty < \lambda < \infty \text{ and } 0 < \alpha < 2\pi, \tag{15}$$

it is easily verified that $d\alpha = 2(1 + \lambda^2)^{-1} d\lambda$. Hence, if S denotes the set of values α on $(0, 2\pi)$ for which $e^{i\alpha}$ is in $\text{sp}(U)$, then by (11),

$$\text{meas sp}(U) = \int_S d\alpha = 2\int_{\text{sp}(A)} (1 + \lambda^2)^{-1} d\lambda \leqq 2 \text{ meas sp}(A). \tag{16}$$

If now t denotes a positive real number and $A_t = tA$, then, since

$$tC = A_t B - BA_t,$$

relations (14) and (16) yield

$$2 \text{ meas sp}(A_t) \geqq 4\pi t \| C \|/d_B(1 + t^2 \| A \|^2). \tag{17}$$

Since meas sp$(A_t) = t$ meas sp(A), division of (17) by t followed by $t \to 0$, leads to the desired relation (4). This completes the proof of (I).

3. Proof of (II). Relations (2) and (3) imply that A is absolutely continuous on $\mathfrak{R}(C)$, that is, that $\| E(\lambda) y \|^2$ is absolutely continuous whenever $y = Cx$; see [10], p. 1027, [11], pp. 514-515, also [12]. As above, relation (9) follows from (2). Let t denote any real number not in the spectrum of J, as defined in (9), and put $J_t = J - tI$. Then $AJ_t - J_tA = -iC$ and hence

$$AJ_t^{-1} - J_t^{-1}A = iJ_t^{-1}CJ_t^{-1} \tag{18}$$

It follows as before that A is absolutely continuous on $\mathfrak{R}(J_t^{-1}CJ_t^{-1})$, that is, on $\mathfrak{R}(J_t^{-1}C)$. Since t is not in sp(J), then also s is not in sp(J) for s sufficiently close to t, hence A is absolutely continuous on $\mathfrak{R}(J_s^{-1}C)$ for $| s - t |$ sufficiently small. Since $\mathfrak{H}_a$ is a subspace, and since $d^n/ds^n (J_s^{-1}) = n! J_s^{-n-1}$, it follows that A is absolutely continuous on the range of $J_t^{-m}C$ for $m = 1, 2, \cdots$, as well as for $m = 0$. Consequently, A is absolutely continuous on the smallest subspace reducing J_t^{-1} and containing $\mathfrak{R}(C)$. But this space is identical with $\mathfrak{M}_B$, the smallest subspace reducing J and containing $\mathfrak{R}(C)$. Since, by (1), the absolute continuity of $\| E(\lambda) x \|^2$ implies that of $\| E(\lambda) A^n x \|^2$ for $n = 0, 1, 2, \cdots$, the assertion (5) follows and the proof of (II) is complete.

4. Applications to integral operators. Throughout this section it will be supposed that $\mathfrak{H} = L^2(a, b)$ where $-\infty < a < b < \infty$. In addition, it will be supposed that the operator A on $L^2(a, b)$ is defined by (8), where $h(t)$ is real-valued, $\phi(t)$ is complex-valued, and where both functions are bounded and measurable on (a, b). If the integral is the Cauchy principal value, then A is a bounded self-adjoint operator on $L^2(a, b)$; cf. Schwartz [13], Section 2, also Dao-xeng [16], p. 244. Henceforth B will denote the multiplication operator on $L^2(a, b)$ defined by

$$Bx = itx(t). \tag{19}$$

Let Z be defined by $Z = \{t \colon a < t < b \text{ and } \phi(t) = 0\}$. It is clear that $L^2(Z)$ reduces both A and B and that for x in $L^2(Z)$, $Ax = h(t)x(t)$. Hence, in order to investigate the operator A, it is sufficient to restrict A to $L^2(Z^*)$, where $Z^* = (a, b) - Z$. There is essentially no loss of generality in supposing that $Z^* = (a, b)$, and there will be proved the following

THEOREM 2. *Let $\phi(t)$ be a complex-valued, bounded, measurable function on (a, b) satisfying*

$$\phi(t) \neq 0 \text{ almost everywhere,} \tag{20}$$

and let $h(t)$ *be an arbitrary real-valued, bounded, measurable function on* (a,b). *Then the bounded self-adjoint operator* A *on* $\mathfrak{H}=L^2(a,b)$ *defined by* (8) *is absolutely continuous, that is,*

$$\mathfrak{H}_a=L^2(a,b), \tag{21}$$

and, in addition,

$$\int_a^b |\phi(t)|^2\,dt \leqq (1/2)(b-a)\operatorname{meas}\operatorname{sp}(A). \tag{22}$$

It is noteworthy that no assumptions of smoothness (not even of continuity) are imposed on either $\phi(t)$ or $h(t)$.

Proof of Theorem 2. It follows from (8) and (19) that

$$(AB-BA)x=\pi^{-1}\int_a^b \phi(t)\phi^*(s)x(s)\,ds\equiv Cx. \tag{23}$$

(Cf. Friedrichs [3], p. 369.) Since $Cx=\pi^{-1}\phi(t)(x,\phi)$, it is clear that $(Cx,x)=\pi^{-1}|(x,\phi)|^2\geqq 0$, so that (3) holds. In addition, it is clear that $\|C\|=\pi^{-1}\|\phi\|^2$. Moreover,

$$(B-B^*)/2i=t, \tag{24}$$

so that $d_B=b-a$ and (22) follows from (4) of Theorem 1.

There remains then to prove (21). It follows from (24) and the definitions of $\mathfrak{M}_B$ that $\mathfrak{M}_B$ contains the closure of the linear manifold on $L^2(a,b)$ spanned by the functions $t^n\phi(t)$, where $n=0,1,2,\cdots$. Relation (20) implies that $\mathfrak{M}_B=L^2(a,b)$, hence $\mathfrak{N}=L^2(a,b)$, and (21) follows from (5) of Theorem 1.

5. A special case. If $h(t)\equiv 0$ and $\phi(t)\equiv 1$, then A of (8) reduces to the operator

$$A_0=(i\pi)^{-1}\int_a^b (s-t)^{-1}x(s)\,ds. \tag{25}$$

According to (22),

$$\operatorname{meas}\operatorname{sp}(A_0)\geqq 2. \tag{26}$$

On the other hand, $\|A\|\leqq 1$; cf. Koppelman and Pincus [6], p. 399, also Dao-xeng [16], Lemma 3, p. 236. Since $\operatorname{meas}\operatorname{sp}(A)\leqq 2\|A\|$, there now follows from Theorem 2 the result that

(27) A_0 is absolutely continuous with the spectrum $[-1,1]$.

(Relation (27), and even more, has also been obtained though the use of

Fourier transforms by Koppelman and Pincus [6]. See also Widom [15], p. 152.)

The above example clearly furnishes an instance of equality in (22) and, in case $b = -a$, also in (6) with $C \geqq 0$ and $C \neq 0$.

6. Remarks. Theorem 2 can be considered as a result concerning certain perturbation of bounded self-adjoint multiplication operators on $L^2(a,b)$. In fact, consider the family T_ϵ defined by

$$T_\epsilon x = h(t)x(t) + \epsilon A_0 x, \qquad \epsilon \geqq 0, \tag{28}$$

and where $h(t)$ is real, bounded, and measurable on (a,b). Then T_0 is the multiplication operator $h(t)$, the spectrum of which, of course, need not be absolutely continuous. Also, since $h(t)$ is not required to be continuous, $\operatorname{sp}(T_0)$ need not be an interval. However, by Theorem 2, whenever $\epsilon > 0$, T_ϵ must be absolutely continuous and $\operatorname{meas}\operatorname{sp}(T_\epsilon) \geqq 2\epsilon$. Also, the spectrum of T_ϵ need not be an interval, even though $\operatorname{sp}(\epsilon A_0)$ is $[-\epsilon, \epsilon]$. In fact, if λ is not in $\operatorname{sp}(T_0)$ then λ is not in $\operatorname{sp}(T_\epsilon)$ for ϵ sufficiently small; hence, if $\operatorname{sp}(T_0)$ is not an interval then neither is $\operatorname{sp}(T_\epsilon)$ for ϵ sufficiently small.

From a different point of view, if the multiplication operator $h(t)$ is regarded as a perturbation of A_0, so that by (8),

$$Tx = h(t)x(t) + A_0 x, \tag{29}$$

then Theorem 2 assures a certain "stability" of the spectrum, both with respect to absolute continuity and to measure. Thus T, as well as A_0, is absolutely continuous and moreover $\operatorname{meas}\operatorname{sp}(T) \geqq 2$ $(= \operatorname{meas}\operatorname{sp}(A_0))$. That, again, $\operatorname{sp}(T)$ need not be an interval, is clear from the observation that $h(t)$ of (29) can be identified with $h(t)/\epsilon$ of (28).

Finally, it can be remarked that Theorem 1 can also be applied to certain singular integral operators similar to (8), but with symmetric kernels $K(t,s)$ more general than that of $\phi(t)\phi^*(s)$, provided that the corresponding operator C (cf. (23)) satisfies the essential restriction (3).

Appendix.

It will be shown below that the assertion of (II) can be clarified in the following equivalent form:

(II′) *Under the assumptions of Theorem* 1,

$$\mathfrak{L} \subset \mathfrak{H}_a, \tag{5′}$$

where $\mathfrak{L}$ *denotes the smallest subspace reducing both operators* A *and* $\operatorname{Im}(B)$, *and which contains the range of* C.

That $\mathfrak{N} \subset \mathfrak{L}$ is clear. It will be shown that $\mathfrak{N} = \mathfrak{L}$, so that (II′) reduces to (II). In order to see this, first note that the space $\mathfrak{N}$ is the closure of the linear manifold of vectors $y = \sum A^m J^n C x_{mn}$ $(m, n \geqq 0)$, where the x_{mn} are in $\mathfrak{H}$ and $J = \operatorname{Im}(B)$. On the other hand, $\mathfrak{L}$ is the closure of the linear manifold spanned by the vectors of the form $z = PCx$, wher x is in $\mathfrak{H}$ and P is a product of a finite number of factors of non-negative powers of A and J. It follows from (9) however that every vector of the type z above can be expressed as a vector of the type y, and hence $\mathfrak{N} = \mathfrak{L}$.

PURDUE UNIVERSITY.

REFERENCES.

[1] T. Carleman, "Sur la résolution de certaines équations intégrales," *Arkiv för Matematik, Astronomie och Fysik*, vol. 16 (1922).

[2] K. O. Friedrichs, "Über die Spektralzerlegung eines Integraloperators," *Mathematische Annalen*, vol. 115 (1938), pp. 249-272.

[3] ———, "Perturbation of continuous spectra," *Communications on Pure and Applied Mathematics*, vol. 1 (1948), pp. 361-405.

[4] T. Kato, "On finite-dimensional perturbations of self-adjoint operators," *Journal of the Mathematical Society of Japan*, vol. 9 (1957), pp. 239-249.

[5] W. Koppelman, "On the spectral theory of singular integral operators," *Transactions of the American Mathematical Society*, vol. 97 (1960), pp. 35-63.

[6] W. Koppelman and J. Pincus, "Spectral representations for finite Hilbert transformations," *Mathematische Zeitschrift*, vol. 71 (1959), pp. 399-407.

[7] S. T. Kuroda, "On the existence and the unitary property of the scattering operator," *Il Nuovo Cimento*, vol. 12 (1959), pp. 431-454.

[8] N. J. Muskhelishvili, *Singular Integral Equations*, Groningen, Noordhoff (1953).

[9] J. D. Pincus, "On the spectral theory of singular integral operators," *to appear.*

[10] C. R. Putnam, "On commutators and Jacobi matrices," *Proceedings of the American Mathematical Society*, vol. 7 (1956), pp. 1026-1030.

[11] ———, "Commutators and absolutely continuous operators," *Transactions of the American Mathematical Society*, vol. 87 (1958), pp. 513-525.

[12] ———, "Commutators, perturbations and unitary spectra," *Acta Mathematica*, vol. 106 (1960), pp. 215-232.

[13] J. Schwartz, "Some non-selfadjoint operators," *Communications on Pure and Applied Mathematics*, vol. 13 (1960), pp. 609-639.

[14] ———, "Some results on the spectra and spectral resolutions of a class of singular integral operators," *Communications on Pure and Applied Mathematics*, vol. 15 (1962), pp. 75-90.

[15] H. Widom, "Singular integral equations in L_p," *Transactions of the American Mathematical Society*, vol. 97 (1960), pp. 131-160.

[16] Xa Dao-xeng, "On non-normal operators," *Chinese Mathematics* (American Mathematical Society translation of *Acta Mathematica Sinica*), vol. 3 (1963), pp. 232-246.

[17] P. R. Halmos, *Introduction to Hilbert Space*, Chelsea, New York (1951).

PARTITIONS INTO ODD AND UNEQUAL PARTS.

By PETER HAGIS, JR.

1. Introduction. In [2] the author has obtained a convergent series and asymptotic formulas for $q(n)$, the number of partitions of a positive integer n into positive odd summands. The necessary transformation equations as well as estimates of the magnitude of certain sums of roots of unity were obtained by using essentially the same procedures as Lehner [4], while the circle dissection method of Rademacher [6] was employed for the integration. In the present paper we shall impose an additional restriction on the partitions, namely that the summands be distinct. That is, we wish to find a convergent series and asymptotic formulas for $Q(n)$, the number of partitions of a positive integer n into odd and unequal parts. $Q(n)$ also represents the number of self-conjugate partitions of n (see pp. 278-9 in [3]). The investigation is parallel to that in [2], and free use will be made of the results obtained in the earlier paper whenever they are applicable.

2. The transformation equations. Several generating functions, each convergent in the interior of the unit circle, will be needed. We list these here.

$$F(x)=\prod_{m=0}^{\infty}(1-x^{2m+1})^{-1}=1+\sum_{n=1}^{\infty}q(n)x^{n}.$$

$$G(x)=\prod_{m=0}^{\infty}(1+x^{2m+1})=F(x)/F(x^{2})=1+\sum_{n=1}^{\infty}Q(n)x^{n}.$$

$$H(x)=\prod_{m=1}^{\infty}(1+x^{m})^{-1}=1/F(x).$$

Now let k be a positive integer, let h satisfy the condition $(h,k)=1$, and let $\Re(z)>0$. Then Theorem 1 in [2] states that

$$F(\exp\{2\pi ih/k-2\pi z/k\})=\omega(h,k)\exp\{\pi(z-1/z)/12k\}\cdot F(\exp\{2\pi ih'/k-2\pi/zk\}) \tag{2.1}$$

if $2\mid k$, and

$$F(\exp\{2\pi ih/k-2\pi z/k\})=2^{-\frac{1}{2}}\chi(h,k)\exp\{\pi(1/z+2z)/24k\}\cdot H(\exp\{2\pi iH'/k-\pi/zk\}) \tag{2.2}$$

if $2\nmid k$.

Received June 28, 1963.

Here, $hh' \equiv -1 \pmod{k}, 2hH' \equiv -1 \pmod{k}$;

$$\omega(h,k) = \exp\{\pi i\sigma(h,k)\}, \tag{2.3}$$
$$\sigma(h,k) = \sum((\mu/k))((h\mu/k)), \quad \mu = 1,3,\cdots,k-1;$$

$$\chi(h,k) = \exp\{\pi i t(h,k)\}, \tag{2.4}$$
$$t(h,k) = \sum((\mu/2k))((h\mu/k)), \quad \mu = 1,3,\cdots,2k-1;$$

$((x)) = x - [x] - \frac{1}{2} + \frac{1}{2}\delta(x)$ where $\delta(x) = 1$ if x is an integer and 0 otherwise.

With the aid of this theorem we can now derive the necessary transformation equations for $G(x)$, the generating function of $Q(n)$. Three cases must be considered.

If $4 \mid k$ and

$$x = \exp\{2\pi ih/k - 2\pi z/k\} \tag{2.5}$$

then $x^2 = \exp\{2\pi ih/k^* - 2\pi z/k^*\}$ where $2k^* = k$ and $2 \mid k^*$. Since $G(x) = F(x)/F(x^2)$ we have, applying (2.1),

$$G(x) = w(h,k)\exp\{\pi(1/z - z)/12k\}G(x'),$$

where

$$x' = \exp\{2\pi ih'/k - 2\pi/zk\}, \tag{2.6}$$

$$w(h,k) = \omega(h,k)/\omega(h,k^*). \tag{2.7}$$

If $k \equiv 2 \pmod{4}$ then $2 \nmid k^*$. Applying (2.1), (2.2) and recalling that $1/H(x) = F(x)$ we have

$$G(x) = 2^{\frac{1}{2}}W(h,k)\exp\{\pi(-z-2/z)/12k\}F(x')F(\exp\{2\pi ig'/k^* - \pi/zk^*\}),$$

where

$$W(h,k) = \omega(h,k)/\chi(h,k^*), \tag{2.8}$$

and $2hg' \equiv -1 \pmod{k^*}$. Since $hh' \equiv -1 \pmod{k^*}$ we have $h' \equiv 2g' \pmod{k^*}$, and it follows that $x' = \exp\{2\pi ig'/k^* - \pi/zk^*\}$. Therefore,

$$G(x) = 2^{\frac{1}{2}}W(h,k)\exp\{\pi(-z-2/z)/12k\}F^2(x').$$

If $2 \nmid k$ then $x^2 = \exp\{4\pi ih/k - 4\pi z/k\}$. From (2.2) we have

$$G(x) = z(h,k)\exp\{\pi(1/2z - 2z)/24k\}H(\exp\{2\pi iH'/k - \pi/zk\})/H(X)$$

where

$$X = \exp\{2\pi iG'/k - \pi/2zk\}, \tag{2.9}$$

$$z(h,k)=\chi(h,k)/\chi(2h,k). \tag{2.10}$$

Here, $4hG'\equiv -1 \pmod{k}$. Since $2hH'\equiv -1 \pmod{k}$ we have $H'\equiv 2G' \pmod{k}$, and it follows that $X^2=\exp\{2\pi i H'/k-\pi/zk\}$. Since $H(X^2)/H(X) = G(X)$ we conclude that

$$G(x)=z(h,k)\exp\{\pi(1/2z-2z)/24k\}G(X).$$

With x, x', X given by (2.5), (2.6), (2.9) respectively we combine these results to obtain

THEOREM 1. *$G(x)$ satisfies the transformation equation*

$$G(x)=w(h,k)\exp\{\pi(1/z-z)/12k\}G(x') \quad \textit{if } 4\mid k,$$
$$G(x)=2^{\frac{1}{2}}W(h,k)\exp\{\pi(-z-2/z)/12k\}F^2(x') \quad \textit{if } k\equiv 2 \pmod{4},$$
$$G(x)=z(h,k)\exp\{1/2z-2z)/24k\}G(X) \quad \textit{if } 2\nmid k.$$

3. Estimates of three exponential sums. In what follows we shall require estimates of the magnitudes of certain exponential sums involving $w(h,k)$, $W(h,k)$, and $z(h,k)$. We first quote, with some minor notational changes, three pertinent results ((3.5), (4.11), (5.7)) from [2] concerning $\omega(h,k)$ and $\chi(h,k)$ as given by (2.3) and (2.4).

If $4\mid k$ then

$$\omega(h,k)=\exp\{2\pi i(\phi(sh+th')/Gk-6\phi(c-2)/G)\}. \tag{3.1}$$

Here, if $3\mid k$ then $f=1$, $G=48$, while if $3\nmid k$ then $f=3$, $G=16$. ϕ is defined by the congruence $f\phi\equiv 1 \pmod{Gk}$, while $hh'\equiv -1 \pmod{Gk}$. $c=1,3$ according as $h\equiv 1,3 \pmod{4}$. Finally,

$$\begin{aligned} s&=2k^2-6k-2,\\ t&=k^2+2. \end{aligned} \tag{3.2}$$

If $k\equiv 2 \pmod{4}$ then

$$\omega(h,k)=\exp\{2\pi i(\lambda(-h+h')/Jk^* + \Gamma(-12d(k+k^*)+4d+6k-3k^2)/e)\}. \tag{3.3}$$

Here, if $3\mid k$ then $e=16$, $J=3$, while if $3\nmid k$ then $e=48$, $J=1$. λ and Γ are defined by the congruences $e\lambda\equiv 1 \pmod{Jk^*}$, $Jk^*\Gamma\equiv 1 \pmod{e}$, while $hh'\equiv -1 \pmod{Jk}$. $d=\pm 3$ according as $h\equiv \pm 3 \pmod{4}$.

If $2\nmid k$ then

$$\chi(h,k)=\exp\{2\pi i(\Phi(-2h-h')/gk+\Upsilon(12k-12k^2+6k\alpha-6)/F)\}. \tag{3.4}$$

Here, if $3\mid k$ then $F=16$, $g=3$, while if $3\nmid k$ then $F=48$, $g=1$. Φ and γ

are defined by the congruences $F\Phi \equiv 1 \pmod{gk}$, $gk\gamma \equiv 1 \pmod{F}$, while $hh' \equiv -1 \pmod{gk}$. α is defined by $k\alpha \equiv 1 \pmod{4}$, $0 < \alpha < 4$.

From these results, similar statements involving $w(h,k)$, $W(h,k)$, and $z(h,k)$ can be derived. We consider four cases.

If $8 \mid k$, then $4 \mid k^*$ and we have from (2.7), (3.1), and (3.2)

$$w(h,k) = \exp\{2\pi i\phi(uh + vh')/Gk\}, \tag{3.5}$$

where $u = k^2 + 2$, $v = 2k^{*2} - 2$.

If $k \equiv 4 \pmod{8}$, then $k^* \equiv 2 \pmod{4}$ and we have from (2.7), (3.1), and (3.3)

$$w(h,k) = \exp\{2\pi i(\phi(sh + th')/Gk - \lambda(-h + h')/JK - 6\phi(c-2)/G \\ - \Gamma(-12d(k^* + K) + 4d + 6k^* - 3k^{*2})/e)\}.$$

Here $K = k^*/2 = k/4$. Since $e = 16f$ and $G = 16J$ the congruences $f\phi \equiv 1 \pmod{Gk}$ and $e\lambda \equiv 1 \pmod{JK}$ imply that $\phi \equiv 16\lambda \pmod{JK}$. If we define r by the congruence $16r \equiv 1 \pmod{JK}$ we see that $\lambda \equiv r\phi \pmod{JK}$. Finally, since $Gk = 64JK$, we have

$$w(h,k) = \exp\{2\pi i(\phi(uh + vh')/Gk - 6\phi(c-2)/G \\ - \Gamma(-12d(k^* + K) + 4d + 6k^* - 3k^{*2})/e)\} \tag{3.6}$$

where $u = 2k^2 - 6k - 2 + 64r$, $v = k^2 + 2 - 64r$.

If $k \equiv 2 \pmod{4}$, then $2 \nmid k^*$ and we have from (2.8), (3.3) and (3.4), noting that $e = F$, $J = g$, $\lambda = \Phi$, $\Gamma = \gamma$,

$$W(h,k) = \exp\{2\pi i(\lambda(2h + 4h')/Jk \\ + \Gamma(-12d(k + k^*) + 4d - 6k^*\alpha + 6)/e)\}. \tag{3.7}$$

If $2 \nmid k$ then from (2.10) and (3.4),

$$z(h,k) = \exp\{2\pi i\Phi(2h - h' + h^*)/gk\}$$

where $2hh^* \equiv -1 \pmod{gk}$. Since $hh' \equiv -1 \pmod{gk}$ we have $h' \equiv 2h^* \pmod{gk}$. If we define j by $2j \equiv 1 \pmod{gk}$ then $h^* \equiv jh' \pmod{gk}$. Therefore,

$$z(h,k) = \exp\{2\pi i\Phi(2h - jh')/gk\}. \tag{3.8}$$

We now state the main results of this section.

THEOREM 2. *The sum* $\sum'_{h \bmod k} w(h,k)\exp\{-2\pi i(hn - h'\nu)/k\}$ *where* $h \equiv d \pmod{4}$, $2 \nmid d$; $a \leqq h' < b \pmod{k}$, $0 \leqq a < b \leqq k$, $4 \mid k$ *is subject to the estimate* $0(n^{1/3}k^{2/3+\epsilon})$ *uniformly in* ν, a, b, d.

THEOREM 3. *The sum* $\sum'_{h \bmod k} W(h,k)\exp\{-2\pi i(hn-h'\nu)/k\}$ *where* $h \equiv d \pmod 4$, $2 \nmid d$; $a \leqq h' < b \pmod k$, $0 \leqq a < b \leqq k$, $k \equiv 2 \pmod 4$, *is subject to the estimate* $0(n^{1/3}k^{2/3+\epsilon})$ *uniformly in* ν, a, b, d.

THEOREM 4. *The sum* $\sum'_{h \bmod k} z(h,k)\exp\{-2\pi i(hn-G'\nu)/k\}$ *where* $a \leqq h' < b \pmod k$, $0 \leqq a < b \leqq k$, $2 \nmid k$ *is subject to the estimate* $0(n^{1/3}k^{2/3+\epsilon})$ *uniformly in* ν, a, b.

Here $hh' \equiv -1 \pmod k$, $4hG' \equiv -1 \pmod k$ and $\sum'$ indicates that h runs over a reduced residue system modulo k. Using (3.5), (3.6), (3.7), and (3.8) the proofs of these theorems are essentially the same as those of Theorems 2 and 3 in [1] and are therefore omitted.

4. A convergent series and asymptotic formulas for $Q(n)$. We now have the tools required for the determination of $Q(n)$. The procedure followed in this section is that of Rademacher, and for the integration details the interested reader is referred to [6]. By Cauchy's formula we have

$$Q(n) = \frac{1}{2\pi i}\int_C x^{-n-1}G(x)\,dx = \frac{1}{2\pi i}\sum_{h,k}'\int_{\xi_{hk}} x^{-n-1}G(x)\,dx$$

where $0 \leqq h < k \leqq N$ and ξ_{hk} are the Farey arcs of order N of C, the circle $|x| = \exp\{-2\pi N^{-2}\}$. If on ξ_{hk} we let $x = \exp\{-2\pi N^{-2} + 2\pi ih/k + 2\pi i\phi\}$ and write $w = N^{-2} - i\phi$, $z = wk$ we obtain

$$Q(n) = \sum_{h,k}'\exp\{-2\pi inh/k\}\int_{-\alpha}^{\beta} G(\exp\{2\pi ih/k - 2\pi z/k\})\exp\{2\pi nw\}\,d\phi.$$

Here $\alpha = 1/k(k+k_1)$ and $\beta = 1/k(k+k_2)$ where $h_1/k_1 < h/k < h_2/k_2$ are consecutive terms in the Farey series of order N.

If we split the sum over k into three parts $Q(n,1)$, $Q(n,2)$, $Q(n,3)$ according as $k \equiv 0, 2, \pm 1 \pmod 4$ respectively, then

$$Q(n) = Q(n,1) + Q(n,2) + Q(n,3).$$

By Theorem 1,

$$Q(n,1) = \sum_{h,k}' w(h,k)\exp\{-2\pi inh/k\}\int_{-\alpha}^{\beta}\sum_{\nu=0}^{\infty} Q(\nu)\exp\{2\pi ih'\nu/k\}$$
$$\cdot\exp\{-(\pi/k^2w)(2\nu - 1/12) + \pi w(2n - 1/12)\}\,d\phi$$

where $4 \mid k$, $k \leqq N$.

Splitting the sum over ν into two parts according as $\nu = 0$ or $\nu > 0$ we find, employing Theorem 2, that

$$Q(n,1) = 2\pi \sum_{k=1}^{N} A(k,n) L(k,n) + 0(n^{1/3} N^{-1/3+\epsilon} \exp\{2\pi n N^{-2}\})$$

where

$$A(k,n) = \sum_{h \bmod k}{}' w(h,k) \exp\{-2\pi i n h/k\}, \tag{4.1}$$

$$L(k,n) = k^{-1}(24n-1)^{-1/2} I_1\{\pi(24n-1)^{1/2}/6k\}. \tag{4.2}$$

Here $I_1(z)$ is the Bessel function of first order and $4 \mid k$.

Turning to $Q(n,2)$ we have by Theorem 1

$$Q(n,2) = 2^{1/2} \sum_{h,k}{}' W(h,k) \exp\{-2\pi i n h/k\} \int_{-\alpha}^{\beta} \sum_{\nu=0}^{\infty} c(\nu) \exp\{2\pi i h'\nu/k\}$$
$$\cdot \exp\{-(\pi/k^2 w)(2\nu + 1/6) + \pi w(2n - 1/12)\} d\phi,$$

where $k \equiv 2 \pmod 4$, $k \leqq N$ and where we have written $F^2(x) = \sum_{\nu=0}^{\infty} c(\nu) x^\nu$. Since $2\nu + 1/6 > 0$ we find, making use of Theorem 3, that

$$Q(n,2) = 0(n^{1/3} N^{-1/3+\epsilon} \exp\{2\pi n N^{-2}\}).$$

Again employing Theorem 1,

$$Q(n,3) = \sum_{h,k}{}' z(h,k) \exp\{-2\pi i n h/k\} \int_{-\alpha}^{\beta} \sum_{\nu=0}^{\infty} Q(\nu) \exp\{2\pi i G'\nu/k\}$$
$$\cdot \exp\{-(\pi/2k^2 w)(\nu - 1/24) + \pi w(2n - 1/12)\} d\phi$$

where $2 \nmid k$, $k \leqq N$.

If we split the sum over ν into two parts according as $\nu = 0$ or $\nu > 0$ we find, making use of Theorem 4, that

$$Q(n,3) = \pi \sum_{k=1}^{N} B(k,n) M(k,n) + 0(n^{1/3} N^{-1/3+\epsilon} \exp\{2\pi n N^{-2}\})$$

where

$$B(k,n) = \sum_{h \bmod k}{}' z(h,k) \exp\{-2\pi i n h/k\}, \tag{4.3}$$

$$M(k,n) = k^{-1}(24n-1)^{-1/2} I_1\{\pi(24n-1)^{1/2}/12k\}, \tag{4.4}$$

and $2 \nmid k$.

Letting $N \to \infty$ we obtain

THEOREM 5. *$Q(n)$, the number of partitions of a positive integer n into distinct odd positive summands is given by the convergent series*

$$Q(n) = 2\pi \sum_{k=1}^{\infty} A(k,n) L(k,n) + \pi \sum_{k=1}^{\infty} B(k,n) M(k,n),$$

where A, L, B, M *are given by* (4.1), (4.2), (4.3), (4.4) *respectively. In the first sum* $4 \mid k$, *in the second sum* $2 \nmid k$.

As in the case of $q(n)$ the dominant term in the series representing $Q(n)$ is that for which $k = 1$. Writing $T = \pi(24n - 1)^{1/2}/12$ we have

THEOREM 6. *As* $n \to \infty$

$$Q(n) = \pi(24n - 1)^{-1/2} I_1\{T\}(1 + O(\exp\{-cn^{1/2}\}))$$

where $c > 0$.

COROLLARY. *As* $n \to \infty$

$$Q(n) = 6^{1/2}(24n - 1)^{-3/4}\exp\{T\}(1 + O(n^{-1/2}))$$

We omit the proofs of these results as they are very similar to those of Theorem 7 and its corollaries in [1].

5. Concluding remarks. The results obtained in this paper can be translated into the language of modular forms. Thus, let $\Gamma_0(4)$ be the subgroup of the modular group consisting of those transformations $V(t) = (at + b)/(ct + d)$ which satisfy the restriction $4 \mid c$. If $c \neq 0$ we can take $c > 0$. If $c = 0$ we have $V(t) = t + b$, with b an arbitrary integer. Thus, $\Gamma_0(4)$ contains the subgroup of translations generated by $S(t) = t + 1$. The circles $|ct + d| = 1$ where $c > 0$, $4 \mid c$, $(c, d) = 1$, are the isometric circles of $\Gamma_0(4)$. It follows that that part of the strip $0 < \mathfrak{R}(t) < 1$ which is exterior to the circles $|4t - 1| = 1$ and $|4t - 3| = 1$ is a fundamental region for $\Gamma_0(4)$ with parabolic cusps at ∞, 0, and 1/2. We also note that $\Gamma_0(4)$ is generated by the transformations $S = t + 1$ and $T = (3t - 1)/(4t - 1)$.

With $k > 0$, $4 \mid k$, $hh' \equiv -1 \pmod{k}$, define $k' = -(hh' + 1)/k$. With $\mathfrak{R}(z) > 0$, but z otherwise arbitrary, $t = h/k + iz/k$ represents a complex variable with domain the upper half-plane. We easily verify that

$$V(t) = (h't + k')/(kt - h) = h'/k + iz/k.$$

If we now define $g(t) = \exp\{-2\pi it/24\} G(\exp\{2\pi it\})$, the first statement of Theorem 1 becomes

$$g(V(t)) = \{w(h,k)\}^{-1}\exp\{2\pi i(h - h')/24k\} g(t),$$

while with $\alpha = 23/24$ we have $g(t + 1) = \exp\{2\pi i\alpha\} g(t)$. Thus, $g(t)$ is a modular form on $\Gamma_0(4)$ of dimension zero with multiplier system ϵ given by

$$\epsilon(h', k', k, -h) = \{w(h,k)\}^{-1}\exp\{2\pi i(h - h')/24k\} \text{ if } k > 0,$$

and

$$\epsilon(1, b, 0, 1) = \exp\{2\pi i b\alpha\}.$$

The second and third statements of Theorem 1 give the behavior of $g(t)$ at the parabolic cusps 1/2 and 0 respectively.

Writing $Q(0) = 1$ we have

$$\exp\{-2\pi i\alpha t\} g(t) = \sum_{m=-1}^{\infty} Q(m+1)\exp\{2\pi i m t\}$$

so that the partition numbers discussed in this paper are seen to be the Fourier coefficients of the modular form $g(t)$. In this setting the referee has pointed out that a weaker form of Theorem 5 is contained as a special case in a theorem due to Lehner (Theorem 2 in [5]). Lehner's result, stated in our notation, is

$$Q(n) = 2\pi \Sigma A(k,n) L(k,n) + \pi \Sigma B(k,n) M(k,n) + O(1).$$

In the first sum $4 \mid k$, $4 \leqq k < \beta(n-1)^{1/2}$. In the second sum $2 \nmid k$, $1 \leqq k < \beta(n-1)^{1/2}$. β is an arbitrary positive constant.

Acknowledgment. The inclusion of the final section of this paper is due to a suggestion by the referee. Both the form and content of this section are largely due to his very pertinent comments. I, of course, take full responsibility for any errors that may be present.

TEMPLE UNIVERSITY, PHILADELPHIA, PENNSYLVANIA.

REFERENCES.

[1] P. Hagis, Jr., "A problem on partitions with a prime modulus $p \geqq 3$," *Transactions of the American Mathematical Society*, vol. 102 (1962), pp. 30-62.

[2] ———, "Partitions into odd summands," *American Journal of Mathematics*, vol. 85 (1963), pp. 213-222.

[3] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford, 1954.

[4] J. Lehner, "A partition function connected with the modulus five," *Duke Mathematical Journal*, vol. 8 (1941), pp. 631-655.

[5] ———, "The Fourier coefficients of automorphic forms on horocyclic groups, II," *Michigan Mathematical Journal*, vol. 6 (1959), pp. 173-193.

[6] H. Rademacher, "The Fourier coefficients of the modular invariant $J(\tau)$," *American Journal of Mathematics*, vol. 60 (1938), pp. 501-512.

PROPERTIES OF CERTAIN TYPES OF WILD SURFACES IN E^3.

By C. E. Burgess.*

1. Introduction. Let S be a wild 2-sphere in E^3 such that $S + \operatorname{Int} S$ is a topological cube, and let W be the set of all wild points on S. In Section 3 we show that if S_1 is a 2-sphere in $S + \operatorname{Int} S$ such that $W \subset S_1$, then S_1 is wild at each point of W. Kinoshita [16] has announced the special case of this result where W is a point and $S_1 - W \subset \operatorname{Int} S$. In Section 4 we show that if W is a subset of a tame Cantor set in E^3, then there is a tame 2-sphere S_2 such that $S \cdot S_2 = W$ and $S_2 \subset S + \operatorname{Ext} S$. Alexander [1] has given an example of S, known as a horned sphere, where the set W is a tame Cantor set, and Ball [3] has described a different such horned sphere. (The construction used by Ball can be changed slightly so that $S + \operatorname{Int} S$ is a topological cube.) Bing [10] recently has described an example of S similar to Alexander's example so that W is a Cantor set in a plane P such that $P \cdot S = W$ and S can be pierced with a straight line interval at each of its points. By using a sphere similar to the one described by Ball, Fort [13] previously gave an example of a wild sphere S that can be pierced by a straight line interval at each of its points. His example can be modified slightly so that W is in a plane P such that $P \cdot S = W$. Fox and Artin [14] have described an example of S where W is a point.

Other examples where W is a subset of a tame Cantor set can be obtained in the following way from the above examples. Let K be a tame 2-sphere in E^3 and let W_0 be a closed subset of a Cantor set on K. Let $D_1, D_2, \cdots$ be a sequence of disjoint disks in $K - W_0$ such that W_0 is the limiting set of this sequence and such that the sequence of diameters of $D_1, D_2, \cdots$ converges to zero. Now for each i, replace D_i with a disk R_i which is wild on a subset W_i of a tame Cantor set in E^3 and is locally tame elsewhere. This can be done so that $(K - \sum_1^\infty D_i) + \sum_1^\infty R_i$ is a sphere S and $S + \operatorname{Int} S$ is a topological cube. The set W of all wild points on S is the set $\sum_0^\infty W_i$. That W is a subset of a tame Cantor set can be seen by applying Bing's theorem [7] that a Cantor

---

Received July 22, 1963.
Revised November 27, 1963.

* This work was supported by the National Science Foundation, under GP-107, and by the Air Force Office of Scientific Research.

set C is tame in E^3 if, for each $\epsilon > 0$, there exist a finite collection of disjoint spheres $K_1, K_2, \cdots, K_n$ such that each K_i has a diameter less than ϵ and $C \subset \sum_1^n \operatorname{Int} K_i$.

In Section 5, we use the results about 2-spheres in Sections 3 and 4 to obtain some similar theorems for connected 2-manifolds in a 3-manifold.

Most of the definitions used here will be similar to those used in [8] and [11]. We use "ϵ-disk" to denote a disk with a diameter less than ϵ. We use "ϵ-neighborhood of a set Y" to denote the set of all points within a distance ϵ of Y. If S is a 2-sphere in E^3, we denote the bounded component and the unbounded component of $E^3 - S$ by "$\operatorname{Int} S$" and "$\operatorname{Ext} S$," respectively. If D is a disk, we use "$\operatorname{Int} D$" to denote $D - \operatorname{Bd} D$.

2. Preliminary theorems.

A simple closed curve is defined to be *unknotted* in E^3 if it is the boundary of a tame disk. We say that an open subset Y of a 2-sphere S can be ***uniformly locally spanned*** in a component V of $E^3 - S$ if for each $\epsilon > 0$, there exists a $\delta > 0$ such that for any δ-disk R in Y and for any positive number α, V contains a disk D with the following properties:

(1) $\operatorname{diam} D < \epsilon$,

(2) there is a homeomorphism of $\operatorname{Bd} R$ onto $\operatorname{Bd} D$ that moves no point more than a distance α.

A complementary domain V of 2-sphere S in E^3 is *locally simply connected* at a point p of S if for each neighborhood N of p there exists a neighborhood U of p such that each map of a simple closed curve into $U \cdot V$ can be shrunk to a point in $N \cdot V$.

THEOREM 1. *If J is a simple closed curve on a 2-sphere S in E^3 and V is a component of $E^3 - S$, then for each positive number ϵ there is an unknotted simple closed curve which is in V and is homeomorphically within ϵ of J.*

Proof. By Bing's Side Approximation Theorem [9], there exists a polyhedral 2-sphere S' and a finite collection $D_1, D_2, \cdots, D_n$ of disjoint $\epsilon/2$-disks on S' such that S' is homeomorphically within $\epsilon/2$ of S and $S' - \sum_1^n D_i \subset V$. There exists a simple closed curve J' such that $J' \subset S'$ and J' is homeomorphically within $\epsilon/2$ of J. Now there is a simple closed curve J'' such that $J'' \subset S' - \sum_1^n D_i$ and J'' is homeomorphically within $\epsilon/2$ of J'. Thus the unknotted simple closed curve J'' is in V and is homeomorphically with ϵ of J.

THEOREM 2. *Suppose S is a 2-sphere in E^3, V is a component of $E^3 - S$, U is an open subset of S, and for each point p of U and each neighborhood N of p there is a neighborhood Q of p such that each unknotted simple closed curve in $Q \cdot V$ is the boundary of a disk in $N \cdot V$. Then V is locally simply connected at each point of U.*

Proof. Let q be a point of U and let H be a disk in U such that $q \in \operatorname{Int} H$. We will show that $\operatorname{Int} H$ can be uniformly locally spanned in V, and it will follow from Theorem 12 of [11] that V is locally simply connected at each point of $\operatorname{Int} H$. (The proof of Theorem 12 of [11] is valid for either complementary domain of S.)

Let ϵ be a positive number. It follows from the hypothesis that there exist open sets $Y_1, Y_2, \cdots, Y_n$ covering H such that for each i $(1 \leqq i \leqq n)$, each unknotted simple closed curve in $V \cdot Y_i$ is the boundary of an ϵ-disk in V. Now let δ be a positive number such that each δ-disk in $\operatorname{Int} H$ is a subset of some Y_i. Let K be a δ-disk in $\operatorname{Int} H$ and let i be an integer so that $K \subset Y_i$. Let α be a positive number. It follows from Theorem 1 that there is an unknotted simple closed curve J in $V \cdot Y_i$ that is homeomorphically with α of $\operatorname{Bd} K$. As J is the boundary of an ϵ-disk in V, we conclude that $\operatorname{Int} H$ can be uniformly locally spanned in V. Thus by Theorem 12 of [11], V is locally simply connected at each point of $\operatorname{Int} H$. This implies that V is locally simply connected at each point of U.

THEOREM 3. *Suppose S and S' are 2-spheres in E^3, R is a disk on S, and D is a disk such that*

(1) $S' \subset S + \operatorname{Int} S$

(2) *R is locally polyhedral at each point of $R - S'$*

(3) $D \subset \operatorname{Ext} S'$

(4) $\operatorname{Bd} D \subset \operatorname{Ext} S$

(5) $D \cdot S \subset R$, and

(6) $3 \operatorname{diam}(D + R) < \operatorname{diam} S'$.

Then for each positive number ϵ there is a disk D' such that

(1) $\operatorname{Bd} D' = \operatorname{Bd} D$

(2) *$D' \subset \operatorname{Ext} S$, and*

(3) *D' is in an ϵ-neighborhood of $D + R$.*

Proof. By Bing's Approximation Theorem [4; 9], we assume that D is locally polyhedral at each point of $\operatorname{Int} D$. Since in any neighborhood of $R \cdot D$, D can be moved slightly to obtain a disk that is in general position with $R - S'$, we assume that D is in general position with $R - S'$; that is, we assume that each component of $D \cdot R$ is a simple closed curve J such that any two points near J, and in different components of $D - J$, are in different components $E^3 - S$. Also, we assume that if X is a disk satisfying all of the requirements in the hypothesis and the above requirements of D such that $\operatorname{Bd} X = \operatorname{Bd} D$ and X is in an ϵ-neighborhood of $D + R$, then the number of components of $X \cdot S$ is not less than the number of components of $D \cdot S$. We wish to show that these assumptions, together with the requirements in the hypothesis, imply that $D \subset \operatorname{Ext} S$. So suppose that D intersects S.

Each component of $D \cdot \operatorname{Int} S$ separates $\operatorname{Int} S$ into two connected sets E and E' such that the boundary of one of them, say of E, is in $D + R$. Now by using the hypotheses that $3 \operatorname{diam} (D + R) < \operatorname{diam} S'$ and $D \subset \operatorname{Ext} S'$, it follows that the closure of E does not intersect S'.

We will show now that no component of $D \cdot (S + \operatorname{Int} S)$ is a disk. Suppose some such component Y is a disk. Then there is a disk W on R such that $W + Y$ is a sphere which, according to the paragraph immediately above, contains no point of S'. Hence W is a polyhedral disk as it is in $R - S'$. Let C be the component of $D - W$ that contains $\operatorname{Bd} D$, and let $L_1, L_2, \cdots, L_n$ denote the simple closed curves in W that are subsets of the boundary of C. We consider these simple closed curves to be ordered so that if $j > i$, then L_j is not in the disk in W that has L_i as its boundary. Now by following a procedure described by Bing [8, p. 297], we can obtain a disk D' which is a subset of $C +$ (an ϵ-neighborhood of W) such that $S \cdot D' \subset D \cdot (S - W)$. This is contrary to the supposition in the first paragraph, as $S \cdot D'$ does not contain $\operatorname{Bd} Y$ and hence $S \cdot D'$ has fewer components than $S \cdot D$. Thus no component of $D \cdot (S + \operatorname{Int} S)$ is a disk.

Now we will show that there is a subdisk D_1 of D such that $D_1 \cdot \operatorname{Int} S$ is a component of $D \cdot \operatorname{Int} S$ and each component of $\operatorname{Int} D_1 - \operatorname{Int} S$ is a disk in $S + \operatorname{Ext} S$. There exists a disk Z_1 in D such that $\operatorname{Int} Z_1$ is a component of $D - S$. It follows from the preceding paragraph that $\operatorname{Int} Z_1$ is a subset of $\operatorname{Ext} S$, so some component of $D \cdot S$ separates Z_1 from $\operatorname{Bd} D$ in D. This implies that there is a disk Y_1 in D such that $\operatorname{Bd} Y_1 \subset R$, $Z_1 \subset \operatorname{Int} Y_1$, and some component X_1 of $Y_1 \cdot \operatorname{Int} S$ has both $\operatorname{Bd} Y_1$ and $\operatorname{Bd} Z_1$ on its boundary. Thus each component of $\operatorname{Int} Y_1 - X_1$ is a disk, so suppose that some component T_1 of $\operatorname{Int} Y_1 - X_1$ is not in $S + \operatorname{Ext} S$. It follows that there is a disk Z_2 in $\operatorname{Int} T_1$ such that $\operatorname{Int} Z_2$ is a component of $T_1 - S$ and $Z_2 \subset S + \operatorname{Ext} S$.

Now there exists a disk Y_2 in T_1 such that $\operatorname{Bd} Y_2 \subset R$, $Z_2 \subset \operatorname{Int} Y_2$, and some component X_2 of $Y_2 \cdot \operatorname{Int} S$ has both $\operatorname{Bd} Y_2$ and $\operatorname{Bd} Z_2$ on its boundary. This procedure can be continued a finite number of steps to obtain a disk D_1 promised in the first sentence of this paragraph.

Let J_1 denote the boundary of D_1, let $J_2, J_3, \cdots, J_m$ denote the components of $S \cdot \operatorname{Int} D_1$, and for each i $(2 \leqq i \leqq m)$, let D_i denote the disk in D_1 that has J_i as its boundary. Also, for each i $(1 \leqq i \leqq m)$, let R_i denote the disk in R that has J_i as its boundary. Let U denote the component of $\operatorname{Int} S - D_1$ that does not intersect S'. There are two cases to be considered, depending upon whether points of $\operatorname{Int} R_1$ near J_1 are on the boundary of U.

Case 1. *Each neighborhood of each point of J_1 contains a point of $\operatorname{Int} R_1$ that is on the boundary of U.* Let K denote the component of $R_1 \cdot \operatorname{Bd} U$ that contains J_1. Since each component of $\operatorname{Bd} K$ is some J_i, assume for convenience that $J_1, J_2, \cdots, J_n$ are the components of $\operatorname{Bd} K$. Let R' denote the polyhedral disk $K + D_2 + D_3 + \cdots + D_n$, and let H denote the component of $D - R'$ that contains $\operatorname{Bd} D$. It follows that $H \subset D - D_1$. Let $L_1, L_2, \cdots, L_r$ denote the simple closed curves in R' that are subsets of the boundary of H, and assume that these simple closed curves are ordered so that if $j > i$, then L_j is not in the disk in R' that has L_i as its boundary. Notice that each L_i is a subset of S. Using the fact that R' is a subset of $R + D \cdot \operatorname{Ext} S$ and does not intersect S', we can apply induction to obtain disjoint simple closed curves $L_1', L_2', \cdots, L_r'$ in D, disjoint disks $H_1, H_2, \cdots, H_r$ in D, and disjoint disks $H_1', H_2', \cdots, H_r'$ in $\operatorname{Ext} S$ such that for each i,

(1) $L_i' \subset \operatorname{Ext} S$,

(2) $L_i + L_i'$ is the boundary of an open annulus in $D \cdot \operatorname{Ext} S$,

(3) $L_i' = \operatorname{Bd} H_i = \operatorname{Bd} H_i'$,

(4) H_i' is in an ϵ-neighborhood of R', and

(5) $(D - \sum_1^r H_i) + \sum_1^r H_i'$ is a disk D'.

Now $S \cdot D' \subset S \cdot D$ and D' is in an ϵ-neighborhood of $D + R$. But D' does not contain J_1, so we have the contradiction that $S \cdot D'$ has fewer components than $S \cdot D$.

Case 2. *Each neighborhood of each point of J_1 contains a point of $S - R_1$ that is on the boundary of U.* From requirement (6) of the hypothesis and the requirement that U contains no point of S', it follows that $\operatorname{Bd} U$ is a subset of $D + R$. Hence for some i $(2 \leqq i \leqq m)$, J_i separates $\operatorname{Bd} R$ from J_1 in R. For convenience, assume that J_2 separates $\operatorname{Bd} R$ from J_1 in R and that for each i $(3 \leqq i \leqq m)$, J_i does not separate J_1 from J_2 in R. It follows that some component K of $R \cdot \operatorname{Bd} U$ contains both J_1 and J_2. For convenience, assume that $J_1, J_2, \cdots, J_n$ are the components of $\operatorname{Bd} K$. Let R' denote the polyhedral disk $K + D_2 + D_3 + \cdots + D_n$. Now $R' \subset R + D \cdot \operatorname{Ext} S$ and does not intersect S'. As in Case 1, we can obtain a disk D' such that $\operatorname{Bd} D' = \operatorname{Bd} D$, D' is in an ϵ-neighborhood of $D + R$, $S \cdot D' \subset S \cdot D$, and D' does not contain J_1. Thus we have the contradiction that $S \cdot D'$ has fewer components than $S \cdot D$.

3. Spheres which intersect at their wild points. A 2-sphere S in E^3 is defined to be *locally tame at a point p from a component V of $E^3 - S$* if there exists a subset U of $S + V$ such that $p \in U$, U is open relative to $S + V$, and the closure of U is a topological cube. Some conditions under which a surface is tame from one of its complementary domains are given in [11].

THEOREM 4. *Suppose S and S' are 2-spheres in E^3, p is a point of $S \cdot S'$, and Y is an open subset of S such that*

(1) $p \in Y$
(2) $S' \subset S + \operatorname{Int} S$
(3) *S' is locally tame at p, and*
(4) *S is locally tame at each point of $Y - S'$.*

Then S is locally tame at p from $\operatorname{Ext} S$.

Proof. There exists a tame disk K' on S' such that $p \in \operatorname{Int} K'$. Let K be a disk on S such that $p \in \operatorname{Int} K$, $K \subset Y$, and $S' \cdot \operatorname{Int} K \subset \operatorname{Int} K'$. We will show that $\operatorname{Ext} S$ is locally simply connected at each point of $\operatorname{Int} K$, and this will imply that S is locally tame from $\operatorname{Ext} S$ at each point of $\operatorname{Int} K$ [11, Theorems 1 and 2].

Let q be a point of $\operatorname{Int} K$ and let N be a neighborhood of q such that $3 \operatorname{diam} N < \operatorname{diam} S'$. For convenience, we assume that K' is a polyhedron and that each of the disks K and K' is a subset of N. Let U be a spherical neighborhood of q such that $U \subset N$, $U \cdot S \subset K$, and $U \cdot S' \subset K'$. Let J be an unknotted simple closed curve in $U \cdot \operatorname{Ext} S$. Thus J is the boundary of a disk D in U. By applying Bing's Approximation Theorem [4; 9] we can assume that D is locally polyhedral at each point of $\operatorname{Int} D$. Furthermore, since K' is a polyhedron and $D \cdot S' \subset K'$, we can follow the procedure described by Bing [8, p. 297] to adjust $\operatorname{Int} D$ so that $D \subset N \cdot \operatorname{Ext} S'$. Also, since S is locally tame at each point of $Y - S'$, we can assume that S is locally polyhedral at each point of $Y - S'$ [18; 5]. Now it follows from Theorem 3 that there is a disk D' such that $\operatorname{Bd} D' = \operatorname{Bd} D$ and $D' \subset N \cdot \operatorname{Ext} S$. Thus we have shown that J is the boundary of a disk in $N \cdot \operatorname{Ext} S$. It follows from Theorem 2 that $\operatorname{Ext} S$ is locally simply connected at each point of $\operatorname{Int} K$.

THEOREM 5. *Suppose S and S' are 2-spheres in E^3, W is the set of all wild points of S, $S + \operatorname{Int} S$ is a topological cube, and $W \subset S' \subset S + \operatorname{Int} S$. Then S' is wild at each point of W.*

Proof. Suppose that S' is locally tame at some point p of W. By Theorem 4 and the hypothesis that $S + \operatorname{Int} S$ is a topological cube, this leads to the contradiction that S is locally tame at p.

4. Spheres which are tame except on a subset of a tame Cantor set.

By considering the theorems of this section and the examples in references cited in Section 1, it can be seen that Theorem 5 does not hold true without the requirement that $S + \operatorname{Int} S$ is a topological cube. Similarly, it follows from Theorem 5 that the theorems of this section do not hold true without this requirement. Theorems 6, 7, and 8 do not hold true without the requirement that W, the set of all wild points of S, is a subset of a tame Cantor set. Antoine [2] has described a 2-sphere S so that W is a wild Cantor set and $S + \operatorname{Int} S$ is a topological cube, and some examples where $W = S$ and $S + \operatorname{Int} S$ is a topological cube have been described by Bing [6] and Martin [17].

We say that a 2-sphere S in E^3 can be *locally spanned at the point p from the component V of $E^3 - S$* if, for each positive number ϵ, there exist disks D and R such that (1) $p \in \operatorname{Int} R \subset S$, (2) $\operatorname{Int} D \subset V$, (3) $\operatorname{Bd} D = \operatorname{Bd} R$, and (4) $\operatorname{diam}(D + R) < \epsilon$. We say that S can be *locally spanned* if it can be locally spanned at each of its points from each component of $E^3 - S$. Such a sphere is tame in E^3 [11].

THEOREM 6. *If S is a 2-sphere in E^3 such that $S + \operatorname{Int} S$ is a topological cube and the set W of all wild points of S is a subset of a tame Cantor set in E^3, then there exists a tame 2-sphere S' such that $S' \subset S + \operatorname{Ext} S$ and $S \cdot S' = W$.*

The following lemma will be used in the proof of Theorem 6.

LEMMA 1. *If S and W satisfy the hypothesis of Theorem 6 such that S is locally polyhedral at each point of $S - W$, then for each $\epsilon > 0$ there exist disjoint polyhedral 2-spheres $K_1, K_2, \cdots, K_n$ with disjoint interiors such that*

(1) $W \subset \sum_1^n \operatorname{Int} K_n$,

(2) *each $K_i \cdot \operatorname{Ext} S$ is connected,*

(3) *each component of each* $K_i \cdot (S + \operatorname{Int} S)$ *is a disk,*

(4) *each* K_i *has a diameter less than* ϵ.

(5) *each* $\operatorname{Int} K_i \cdot \operatorname{Ext} S$ *is connected, and*

(6) *each* $\operatorname{Int} K_i$ *intersects* W.

Proof of Lemma 1. Let δ be a positive number such that $\delta < \epsilon/3$ and each subset of S of diameter less than δ lies in an $\epsilon/3$ disk on S. Since W is a subset of a tame Cantor set, there exist disjoint polyhedral 2-spheres $H_1, H_2, \cdots, H_m$ with disjoint interiors such that each H_i has a diameter less than δ and $W \subset \sum_1^m \operatorname{Int} H_i$ [7]. We assume that each H_i is in general position with $S - W$. Each $H_i \cdot \operatorname{Ext} S$ has only a finite number of components, so let $E_1, E_2, \cdots, E_r$ denote the components of $\operatorname{Ext} S \cdot \sum_1^m H_i$. For each i $(1 \leqq i \leqq r)$, the boundary of E_i is a subset of some $\epsilon/3$-disk D_i in S and is the sum of a finite number of simple closed curves in $S - W$. Using the hypothesis that $S + \operatorname{Int} S$ is a topological cube, we can follow the procedure described by Bing [8, p. 297] to obtain, for each i $(1 \leqq i \leqq r)$, disjoint polyhedral disks $D_{i1}, D_{i2}, \cdots, D_{ir_i}$ such that (1) each $\operatorname{Int} D_{ij}$ is a subset of $\operatorname{Int} S$, (2) each D_{ij} is in an $\epsilon/3$-neighborhood of D_i, and (3) the set $E_i + D_{i1} + D_{i2} + \cdots D_{ir_i}$ is a polyhedral 2-sphere K_i. The above procedure can be followed so that $K_1, K_2, \cdots, K_r$ are disjoint. For convenience, assume that each $\operatorname{Int} K_i$ intersects W.

Now we will show that $W \subset \sum_1^r \operatorname{Int} K_i$. Suppose that some point p of W is not in $\sum_1^r \operatorname{Int} K_i$. Let q be a point in $\operatorname{Ext} S$ that is in neither of the sets $\sum_1^r (K_i + \operatorname{Int} K_i)$ and $\sum_1^m (H_i + \operatorname{Int} H_i)$. There is an arc pq such that $pq - p \subset \operatorname{Ext} S$, pq is locally polyhedral at each point of $pq - q$, and $pq \cdot \sum_1^r K_i$ is finite. Furthermore assume that no other such arc from p to q intersects $\sum_1^r K_i$ in fewer points than does pq. As $p \in \sum_1^m \operatorname{Int} H_i$, it follows that pq intersects one of the sets $E_1, E_2, \cdots, E_r$. Consider the case where pq intersects E_1. Both p and q are in $\operatorname{Ext} K_1$, so let X and Y be the components of $pq \cdot \operatorname{Ext} K_1$ that contain p and q, respectively. Let x and y be the points of E_1 that are limit points of X and Y, respectively. Since E_1 is connected, there is an arc Z in E_1 such that $x + y \subset Z$. Let U be a connected open set which contains Z and does not intersect $S + \sum_2^r K_i$. There is an arc A from p to q such that $A \subset X + Y + U$, $A \subset \operatorname{Ext} K_1$, and A is locally polyhedral except at p. This involves the contradiction that there are fewer points in $A \cdot \sum_1^r K_i$ than in $pq \cdot \sum_1^r K_i$. Hence we have shown that $W \subset \sum_1^r \operatorname{Int} K_i$.

As at least one of the 2-spheres $K_1, K_2, \cdots, K_r$ is not a subset of $\sum_1^r \operatorname{Int} K_i$, assume for convenience that no one of the spheres $K_1, K_2, \cdots, K_n$ is a subset of $\sum_1^r \operatorname{Int} K_i$ and that each of the spheres $K_{n+1}, K_{n+2}, \cdots, K_r$ is a subset of $\sum_1^r \operatorname{Int} K_i$. It follows that the spheres $K_1, K_2, \cdots, K_n$ satisfy the requirements in the conclusion of Lemma 1.

Proof of Theorem 6. If T is a polyhedral annulus and the simple closed curves J and J' are the components of the boundary of T, we call T a *tube* from J to J'.

Let S_0 be a polyhedral 2-sphere such that $S \subset \operatorname{Int} S_0$. Let $K_{11}, K_{12}, \cdots, K_{1n_1}$ be polyhedral 2-spheres which satisfy the requirements of the conclusion of Lemma 1, where $\epsilon = 1$, such that each $K_{1j} \subset \operatorname{Int} S_0$. Let $D_{11}, D_{12}, \cdots, D_{1n_1}$ be disjoint polyhedral disks on S_0 and for each j $(1 \leqq j \leqq n_1)$, let R_{1j} be a polyhedral disk in $\operatorname{Int} K_{1j} \cdot \operatorname{Ext} S$. For each j $(1 \leqq j \leqq n_1)$, there exists a tube T_{1j} from $\operatorname{Bd} D_{1j}$ to $\operatorname{Bd} R_{1j}$ such that

(1) $T_{1j} - \operatorname{Bd} D_{1j} \subset \operatorname{Int} S_0 \cdot \operatorname{Ext} S$,
(2) $T_{1j} \cdot K_{1j}$ is a simple closed curve L_{1j},
(3) T_{1j} does not intersect $\operatorname{Int} R_{1j}$, and
(4) T_{1j} does not intersect K_{1r}, where $r \neq j$.

Furthermore, we require that these tubes be disjoint. Let S_1 denote the polyhedral 2-sphere

$$(S_0 - \sum D_{1j}) + \sum T_{1j} + \sum R_{1j}.$$

Notice that for each j $(1 \leqq j \leqq n_1)$, $S_1 \cdot K_{1j} = L_{1j}$.

Now let $K_{21}, K_{22}, \cdots, K_{2n_2}$ be polyhedral 2-spheres which satisfy the requirements of the conclusion of Lemma 1, where $\epsilon = 1/2$, such that each K_{2j} is in $\operatorname{Int} S_1$ and is in some $\operatorname{Int} K_{1i}$. There exist disjoint polyhedral disks $D_{21}, D_{22}, \cdots, D_{2n_2}$, disjoint polyhedral disks $R_{21}, R_{22}, \cdots, R_{2n_2}$, and disjoint tubes $T_{21}, T_{22}, \cdots, T_{2n_2}$ such that for each j $(1 \leqq j \leqq n_2)$,

(1) D_{2j} is a subset of some R_{1i},
(2) $R_{2j} \subset \operatorname{Int} K_{2j} \cdot \operatorname{Ext} S$,

(3) T_{2j} is a tube from $\operatorname{Bd} D_{2j}$ to $\operatorname{Bd} R_{2j}$,

(4) T_{2j} is in some $\operatorname{Int} K_{1i}$,

(5) T_{2j} does not intersect $\operatorname{Int} R_{2j}$,

(6) $T_{2j} - \operatorname{Bd} D_{2j} \subset \operatorname{Int} S_1 \cdot \operatorname{Ext} S$,

(7) T_{2j} does not intersect K_{2r}, where $r \neq j$, and

(8) $T_{2j} \cdot K_{2j}$ is a simple closed curve L_{2j}.

Let S_2 denote the polyhedral 2-sphere

$$(S_1 - \sum D_{2j}) + \sum T_{2j} + \sum R_{2j}.$$

By use of induction, this process can be continued to obtain a sequence of polyhedral 2-spheres $S_0, S_1, S_2, \cdots$ that converges to a 2-sphere S' such that

(1) $S' \cdot S = W$,

(2) $S' \subset S + \operatorname{Ext} S$,

(3) for each $i \geqq 1$, $S' \supset S_{i-1} - (D_{i1} + D_{i2} + \cdots + D_{in_i})$, and

(4) each K_{ij} has a diameter less than $1/i$.

Thus S' is locally polyhedral at each point of $S' - W$ and each $K_{ij} \cdot S'$ is a simple closed curve L_{ij}.

It remains to be shown that S' is tame, and in particular, that S' is locally tame at each point of W. Let p be a point of W and let δ be a positive number. There exists an integer i such that $1/i < \delta$ and there exists an integer j such that $p \in \operatorname{Int} K_{ij}$. Since $K_{ij} \cdot S' = L_{ij}$, it follows that S' can be locally spanned at p from each component of $E^3 - S'$. Thus S' can be locally spanned at each of its points. It follows from [11, Theorem 8] that S' is tame.

THEOREM 7. *If S is a 2-sphere in E^3 such that $S + \operatorname{Int} S$ is a topological cube and the set W of all wild points of S is a subset of a tame arc in E^3, then there exists a tame 2-sphere S' such that $S' \subset S + \operatorname{Ext} S$ and $S \cdot S' = W$.*

Proof. Let H be a tame arc in E^3 such that $W \subset H$. Suppose that W is not a subset of a Cantor set in H. Then some open interval L in H is a subset of W. Let U be an open subset of S that intersects L and contains no point of $H - L$. Then S is locally tame at each point of $U - L$, and there is a tame arc K in $L \cdot U$. It follows from a theorem proved by Doyle and Hocking [12] that S is locally tame at each point of K. But this is a contradiction as $K \subset W$. Hence W is a subset of a tame Cantor set in H. The conclusion of Theorem 7 now follows from Theorem 6.

THEOREM 8. *If S is a 2-sphere in E^3 such that $S + \operatorname{Int} S$ is a topological cube and the set W of all wild points of S is a subset of either a tame Cantor set or a tame arc in E^3, then there exists a plane P in E^3 and a homeomorphism h of E^3 onto itself such that $P \cdot h(S) = h(W)$ and $h(W)$ is a subset of a straight line in P.*

Proof. As shown in the proof of Theorem 7, W is a subset of a tame Cantor set in E^3. Follow the procedure in the proof of Theorem 6, where the tame sphere S_0 is replaced with a plane P_0 that does not intersect S, to obtain a tame topological plane P' such that $S \cdot P' = W$. Let h_1 be a homeomorphism of E^3 onto itself that carries P' onto a plane P. As shown by Moore and Kline [19], $h_1(W)$ lies in an arc H in P. (Also, see Moore's remark [20, p. 380] about an earlier proof by Shoenflies.) There is a homeomorphism h_2 of E_3 onto itself such that $h_2(P) = P$ and $h_2(H)$ is a subset of a straight line in P. Then $h_2 h_1$ is a homeomorphism satisfying the requirements in the conclusion of Theorem 8.

5. Generalizations to 2-manifolds in a 3-manifold. Let M be a connected 2-manifold in a triangulated 3-manifold M^3 such that $M^3 - M$ is the sum of two connected sets V_1 and V_2. The manifold M can be defined to be locally tame from V_1 at a point p of M by using the definition given for a 2-sphere in Section 3. We say that M is tame from V_1 if M is locally tame from V_1 at each point of M. This is equivalent to saying that $M + V_1$ is a 3-manifold with boundary.

THEOREM 9. *Suppose M and M' are connected 2-manifolds in a triangulated 3-manifold M^3, p is a point of $M \cdot M'$, and Y is an open subset of M such that*

(1) $p \in Y$,
(2) *M' is locally tame at p,*
(3) *M is locally tame at each point of $Y - M'$,*
(4) *$M^3 - M$ is the sum of two connected sets V_1 and V_2,*
(5) *$M^3 - M'$ is the sum of two connected sets, and*
(6) $M' \subset V_1 + M$.

Then M is locally tame at p from V_2.

Proof. As the connected set V_2 does not intersect M', it follows that there is a component V of $M^3 - M'$ such that $V \subset V_1$ and $M' + V \subset M + V_1$.

Let N be a neighborhood of p that is homeomorphic with E^3, and consider

N to have a triangulation derived from the given triangulation of M^3. By a modification [11, Theorem 1] of a theorem proved by Bing [8, Theorem 5], there exist a disk D and a 2-sphere S such that

(1) $S \subset N$,
(2) $p \in \operatorname{Int} D \subset Y \cdot S$, and
(3) p is on the boundary of some component X of $V_1 \cdot \operatorname{Int} S$, where $\operatorname{Int} S$ denotes the component of $M^3 - S$ that is a subset of N.

Let U be an open set in N that contains p such that $U \cdot V_1 = U \cdot \operatorname{Int} S = U \cdot X$ and $U \cdot M = U \cdot S = U \cdot \operatorname{Int} D$. There exists a disk D' in $U \cdot M'$ such that $p \in \operatorname{Int} D'$ and M' is locally tame at each point of D'. Thus there exists a disk D'' in U such that $\operatorname{Int} D'' \subset V$ and $\operatorname{Bd} D'' = \operatorname{Bd} D'$. Let S' denote the 2-sphere $D' + D''$. Then $p \in S'$, $S' \subset M + V_1$, and $S' \cdot S = D' \cdot D$. As $\operatorname{Int} D'' \subset U \cdot V \subset U \cdot V_1 = U \cdot \operatorname{Int} S$, it follows that $S' \subset S + \operatorname{Int} S$. Thus S and S' satisfy the requirements of the hypothesis of Theorem 4, where E^3 is replaced with N. Hence by Theorem 4, S is locally tame at p from $N - (S + \operatorname{Int} S)$. This implies that M is locally tame at p from V_2.

THEOREM 10. *Suppose M and M' are connected 2-manifolds in a triangulated 3-manifold M^3 such that*

(1) *$M^3 - M$ is the sum of two connected sets V_1 and V_2,*
(2) *M is tame from V_1,*
(3) *$M^3 - M'$ is the sum of two connected sets,*
(4) *$M' \subset M + V_1$, and*
(5) *$W \subset M \cdot M'$, where W is the set of all wild points of M.*

Then M' is wild at each point of W.

Proof. Suppose that M' is locally tame at some point p of W. By Theorem 9 and requirement (2) of the hypothesis, this leads to the contradiction that M is locally tame at p.

The following theorem, which is a direct corollary of Theorem 10, is similar to a theorem proved by Hosay [15]. His requirement that the set of all diameters of components of $M \cdot M'$ has a positive lower bound is omitted, but we add the requirement that M is tame from V_1. That it is necessary to add this requirement can be seen by applying Theorem 6 to some examples in references cited in Section 1.

THEOREM 11. *Suppose M and M' are connected 2-manifolds in a triangulated 3-manifold M^3 such that*

(1) *$M^3 - M$ is the sum of two connected set V_1 and V_2,*
(2) *M is tame from V_1,*
(3) *$M^3 - M'$ is the sum of two connected sets,*
(4) *M is locally tame at each point of $M - M'$,*
(5) *$M' \subset M + V_1$, and*
(6) *M' is tame.*

Then M is tame.

We define a Cantor set C to be *tame* in a 3-manifold M^3 if there is a homeomorphism of M^3 onto itself that carries C into a polygonal arc.

THEOREM 12. *Suppose M is a connected 2-manifold in a triangulated 3-manifold M^3, $M^3 - M$ is the sum of two connected sets V_1 and V_2, M is tame from V_1, and the set W of all wild points of M is a subset of either a tame arc or a tame Cantor set in M^3. Then there is a tame 2-sphere S such that $S \subset M + V_2$ and $S \cdot M = W$.*

Proof. As in the proof of Theorem 7, W is a subset of a tame Cantor set in M^3. Let S_0 be a tame 2-sphere in V_2. The procedure in the proof of Theorem 6 can be followed to obtain a tame 2-sphere S such that $S \subset M + V_2$ and $S \cdot M = W$.

THE INSTITUTE FOR ADVANCED STUDY
AND
THE UNIVERSITY OF UTAH.

REFERENCES.

[1] J. W. Alexander, "An example of a simply connected surface bounding a region which is not simply connected," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 10 (1924), pp. 8-10.

[2] L. Antoine, "Sur l'homéomorphie de deux figures et leurs voisinages," *Journal de Mathématiques Pures et Appliquées*, vol. 4 (1921), pp. 221-325.

[3] B. J. Ball, "The sum of two solid horned spheres," *Annals of Mathematics*, vol. 69 (1959), pp. 253-257.

[4] R. H. Bing, "Approximating surfaces with polyhedral ones," *Annals of Mathematics*, vol. 65 (1957), pp. 456-483.

[5] ———, "An alternative proof that 3-manifolds can be triangulated," *Annals of Mathematics*, vol. 69 (1959), pp. 37-65.

[6] ———, "A wild surface each of whose arcs is tame," *Duke Mathematical Journal*, vol. 28 (1961), pp. 1-16.

[7] ———, "Tame Cantor sets in E^3," *Pacific Journal of Mathematics*, vol. 11 (1961), pp. 435-446.

[8] ———. "A surface is tame if its complement is 1-ULC," *Transactions of the American Mathematical Society*, vol. 101 (1961), pp. 294-305.

[9] ———, "Approximating surfaces from the side," *Annals of Mathematics*, vol. 77 (1963), pp. 145-192.

[10] ———, "Spheres in E^3," *The American Mathematical Monthy* (to appear).

[11] C. E. Burgess, "Characterizations of tame surfaces in E^3," *Transactions of the American Mathematical Society* (to appear).

[12] P. H. Doyle and J. H. Hocking, "Some results on tame disks and spheres in E^3," *Proceedings of the American Mathematical Society*, vol. 11 (1960), pp. 832-836.

[13] M. K. Fort, Jr., "A wild sphere which can be pierced at each point by a straight line segment," *Proceedings of the American Mathematical Society*, vol. 14 (1963), pp. 994-995.

[14] R. H. Fox and E. Artin, "Some wild cells and spheres in three-dimensional space," *Annals of Mathematics*, vol. 49 (1948), pp. 979-990.

[15] Norman Hosay, "Conditions for tameness of a 2-sphere which is locally tame modulo a tame set," *American Mathematical Society Notices*, vol. 9 (1962), p. 117.

[16] S. Kinoshita, "On quasi translations in 3-space," *Topology of 3-Manifolds and Related Topics*, Prentice Hall, 1962, pp. 223-226.

[17] Joseph Martin, "A rigid sphere," *American Mathematical Society Notices*, vol. 10 (1963), p. 189.

[18] E. E. Moise, "Affine structures in 3-manifolds, V. The triangulation theorem and Hauptvermutung," *Annals of Mathematics*, vol. 56 (1952), pp. 96-114.

[19] R. L. Moore and J. R. Kline, "On the most general plane closed point-set through which it is possible to pass a simple continuous arc," *Annals of Mathematics*, vol. 20 (1919), pp. 218-223.

[20] R. L. Moore, *Foundations of Point Set Theory*, American Mathematical Society Colloquium Publications, vol. 13, Revised Edition, 1962.

STRONGLY NON-LINEAR PARABOLIC BOUNDARY VALUE PROBLEMS.

By FELIX E. BROWDER.[1]

In a number of recent papers, the writer has developed a new non-linear version of the orthogonal projection argument to prove the existence of solutions of variational type for boundary value problems for non-linear partial differential equations, first for non-linear elliptic equations in [2], [3], [4], [5] and [6] and more recently for non-linear parabolic equations. This new method is based on general theorems assuring the existence of solutions of non-linear functional equations in Hilbert spaces and reflexive Banach spaces involving operators satisfying very weak continuity conditions but having suitable monotonicity properties.[2] Our results in [6] included much stronger versions of results of M. I. Visik ([12], [13], [14]) on solutions of strongly elliptic boundary value problems for equations of the form

$$Au = \sum_{|\alpha|\leq m} D^{\alpha}uA_{\alpha}(x, y, u, \cdot\cdot, D^{m}u) = f(x), \tag{1}$$

with A_α having polynomial growth in $(u, \cdots, D^m u)$.[3] In the parabolic case, we treat mixed initial-boundary value problems for equations of the form

$$\frac{\partial u}{\partial t} + \sum_{|\alpha|\leq m} D^{\alpha}A_{\alpha}(x, t, u, \cdots, D^{m}u) = f(x, t). \tag{2}$$

It is our purpose in the present paper to give a complete treatment of the parabolic problems for the case where the A_α have polynomial growth in $(u, \cdots, D^m u)$. This treatment is based on a new and sharper theorem on non-linear functional equations in reflexive Banach spaces, Theorem 2 below, which contains as special cases the abstract theorem established in [6].[4]

Received August 14, 1963.
Revise received January 4, 1964.

[1] The preparation of this paper was partially supported by N. S. F. Grant-19751. The author is a Sloan Fellow.

[2] Earlier but weaker results on continuous monotone functional equations were obtained by Vainberg and Kachurovski [11] and G. J. Minty [10].

[3] For the case where the A_α have linear growth and A is the set of Euler equations of a regular variation problem, an interesting generalization of Morse theory applying to solutions of the system (1) has been obtained recently by S. Smale and R. Palais.

[4] After the preparation of the present paper was completed, the writer has learned

The principal results of this paper have been announced in [8].[5]

Section 1 presents the formulation of our non-linear parabolic boundary value problems of variational type and states our main existence theorem, Theorem 1. In Section 2, we state and prove the basic abstract result, Theorem 2. In Section 3, we prove Theorem 1 using Theorem 2.

Section 1. Let Ω be an open subset of the Euclidean space R^n. The points of Ω will be denoted by $x=(x_1,\cdots,x_n)$ and derivatives with respect to the x-variables by

$$D_j=\frac{1}{i}\frac{\partial}{\partial x_j};\ 1\leqq j\leqq n,$$

$$D^\alpha=D_1^{\alpha_1}\cdots D_n^{\alpha_n}\text{ if }\alpha=(\alpha_1,\cdots,\alpha_n),$$

$$|\alpha|=\sum_{j=1}^{n}\alpha_j.$$

The points of R^1 will be denoted by t and differentiation in t by $\frac{\partial}{\partial t}$.

We shall write integrals in Lebesgue n-measure over Ω as $\int f(x)\,dx$.

By functions, we shall mean r-vector functions $u=(u_1,\cdots,u_r)$ for some fixed $r\geqq 1$, where each u_j is a complex valued function on Ω or on $\Omega\times S$ where $S\subset R^1$. Thus $D^\alpha u=(D^\alpha u_1,\cdots,D^\alpha u_r)$, $\frac{\partial u}{\partial t}=\left(\frac{\partial u_1}{\partial t},\cdots,\frac{\partial u_r}{\partial t}\right)$. If u and v are two functions on Ω, we shall denote by $\langle u,v\rangle$ their inner product in $L^2(\Omega)$ i.e.

$$\langle u,v\rangle=\sum_{j=1}\int u_j(x)\bar{v}_j(x)\,dx,$$

(even when neither function lies in $L^2(\Omega)$, as long as the integrals converge.)

Let $S=[t_0,t_1]$ be a closed line segment in R^1. Let $A(t)$ be a system of partial differential operators on Ω with coefficients depending on t in R^1, of the form

$$(1.1)\qquad A(t)u=\sum_{|\alpha|\leqq m}D^\alpha A_\alpha(x,t,u,\cdots,D^m u)$$

---

of papers of Visik, Dokladi A. N. S. S. S. R., vol. 140 (1960), pp. 998-1001 and Mat. Sbornik, vol. 59 (1962), pp. 289-335, which extend the results of [12], [13], and [14] to parabolic initial value problems with lateral Dirichlet boundary conditions. As in the elliptic case, Visik's methods are more concrete-analytic than those of the present paper. A detailed account of his elliptic results under weaker hypotheses than those of [12], [13], and [14] has been published by Visik in *Trudi Mooovskogo Matematičeskogo Obščestva*, vol. 12 (1963), pp. 125-184.

[5] The present paper is an amalgamation of the two papers listed as [7] and [8] in the bibliography of [8].

where each A_α is an r-vector function of (x, t) in $\Omega \times R^1$, of the value of the r-vector function u at (x, t), and the values of all the derivations $D^\beta u$ for $|\beta| \leqq m$. We assume the following about $A(t)$:

ASSUMPTION (I). *Each A_α is measurable in all variables and continuous in $(u, \cdots, D^m u)$.*

For a fixed $p > 1$, there exists a constant $C > 0$ and a function $g \in L^q(\Omega \times R^1)$, where $q = p(p-1)^{-1}$, such that for all (x, t) in $\Omega \times R^1$ and each complex vector $\zeta = \{\zeta_\beta; |\beta| \leqq m\}$

$$|A_\alpha(x, t, \zeta)| \leqq C \sum_{|\beta| \leqq m} |\zeta_\beta|^{p-1} + g(x, t).$$

Let p be the exponent of Assumption (I), $W^{m,p}(\Omega)$ the Banach space defined by

$$W^{m,p}(\Omega) = \{u \mid u \in L^p(\Omega), D^\beta u \in L^p(\Omega) \text{ for } |\beta| \leqq m\},$$

with norm

$$\| u \|_{m,p} = \{ \sum_{|\beta| \leqq m} \| D^\beta u \|^p_{L^p(\Omega)} \}^{1/p}$$

$W^{m,p}(\Omega)$ is a reflexive separable Banach space, since it has a natural embedding as a closed subspace of the direct sum of a finite number of copies of $L^p(\Omega)$.

Let V be a closed subspace of $W^{m,p}(\Omega)$ with $C_c^\infty(\Omega) \subset V$, where $C_c^\infty(\Omega)$ is the family of infinitely differentiable functions with compact support in Ω. We consider the Banach space $L^p(S, V)$ of equivalence classes of functions u from S to V with $\int_{t_0}^{t_1} \| u(t) \|^p_{m,p} \, dt < +\infty$. The norm in this Banach space is

$$\| u \| = \{ \int_{t_0}^{t_1} \| u(t) \|^p_{m,p} \, dt \}^{1/p}$$

$L^p(S, V)$ is itself a reflexive separable Banach space.

It follows from Assumption (I) on the system $A(t)$ that for every u and v in $L^p(S, V)$, the non-linear Dirichlet form $h(u, v)$ is well-defined by:

Definition. For $u, v \in L^p(S, V)$,

$$h(u, v) = \sum_{|\alpha| \leqq m} \int_{t_0}^{t_1} \langle A_\alpha(x, t, u(t), \cdots, D^m u(t)), D^\alpha v(t) \rangle \, dt. \tag{1.2}$$

We know moreover from Assumption (I) that for a suitable constant $c > 0$

$$|h(u, v)| \leqq c \| v \| \cdot \{ \| u \| + 1 \}^{p-1} \tag{1.3}$$

for all $u, v \in L^p(S, V)$.

Definition. *Let*

$F = \{v \mid v \in L^p(S,V)$, *$v$ is continuous from S to V, piecewise continuously differentiable from S to $L^q(\Omega)$*$\}$

$$F_0 = \{v \mid v \in F, v(t_0) = 0\}$$
$$F_1 = \{v \mid v \in F, v(t_1) = 0\}.$$

If $l(u,v)$ is the sequilinear form defined for u in F_0, v in $L^p(S,V)$ by

$$l(u,v) = \int_{t_0}^{t_1} \langle \frac{du}{dt}(t), v(t) \rangle \, dt \tag{1.4}$$

then for all v in $L^p(S,V)$ and each u in F_0,

$$|\, l(u,v) \,| \leqq C(u) \| v \|.$$

i. e. l is a bounded conjugate linear functional of V.

Let $(L^p(S,V))^*$ be the conjugate space of $L^p(S,V)$, i. e. the space of bounded conjugate-linear functionals on $L^p(S,V)$. For u in $L^p(S,V)$, w in $(L^p(S,V))^*$, we denote the pairing of w with u by (w,u). This latter form is therefore linear in w, conjugate-linear in u.

Definition. *δ_0 is the linear mapping of F_0 into $(L^p(S,V))^*$ such that for each u in F_0, $\delta_0 u$ is the element of $(L^p(S,V))^*$ for which all v in $L^p(S,V)$*

$$(\delta_0 u, v) = l(u,v) = \int_{t_0}^{t_1} \langle \frac{du}{dt}(t), v(t) \rangle \, dt. \tag{1.5}$$

Similarly δ_1 is the linear map of F_1 into $(L^p(S,V))^*$ such that for each u in F_1, $\delta_1 u$ is the element of $(L^p(S,V))^*$ for which for all v in $L^p(S,V)$,

$$(\delta_1 u, v) = -\int_{t_0}^{t_1} \langle \frac{du}{dt}(t), v(t) \rangle \, dt. \tag{1.6}$$

We consider both δ_0 and δ_1 as operators with dense domains in $L^p(S,V)$ and ranges in $(L^p(S,V))^*$.

LEMMA 1.1. $\delta_1 \subseteq (\delta_0)^*$.

Proof. Integration by parts.

Definition. *δ is the closure of δ_0 as a linear operator with domain in $L^p(S,V)$ and range in $(L^p(S,V))^*$.*

We know that δ_0 is closable and hence δ is well-defined since by Lemma (1.1), δ_0 has a densely defined adjoint.

Let us now formulate the boundary-value problem which we shall solve, first in a heuristic classical analytic form and then in the precise form of the variational boundary value problem.

Classical parabolic boundary value problem: Given f on $\Omega \times S$, to find u in $C^{2m}(\Omega \times S)$ such that

$$\frac{\partial u}{\partial t} + A(t)u = f(x,t),\ t > t_0$$

$$u(t) \in V,\ t > t_0$$

$$\langle A(t)u(t), v\rangle = \sum_{|\alpha|\leqq m} \langle A_\alpha(x,t,u(t),\cdots,D^m u(t)), D^\alpha v\rangle \quad \text{for all } v \in V,\ t > 0. \tag{1.7}$$

$$u(t_0) = 0.$$

We observed that every solution of the classical parabolic boundary value problem is also a solution of the following variational parabolic boundary problem:

Variational parabolic boundary problem: Given f in $(L^p(S,V))^*$, to find u in $L^p(S,V)$ with $u \in D(\delta)$ such that

$$(\delta u, v) + h(u,v) = (f,v) \tag{1.8}$$

for all v in $L^p(S,V)$.

Our basic result is the following:

THEOREM 1. *Suppose $A(t)$ satisfies Assumption (I) and both of the following conditions are satisfied:*

(1.9)

a) *For all u and v of $L^p(S,V)$,*

$$\operatorname{Re}\{h(u,u-v) - h(v,u-v)\} \geqq 0$$

b) *There exist a non-negative function $c(r)$ with $c(r) \to +\infty$ as $r \to +\infty$ such that*

$$\operatorname{Re} h(u,v) \geqq c(\|u\|)\|u\| \tag{1.10}$$

for all u in $L^p(S,V)$.

Then there exists one and only one solution of the variational boundary value problem for the parabolic system

$$\frac{\partial u}{\partial t} + A(t)u = f$$

for given f in $(L^p(S,V))^*$. This solution u is continuous from S to $L^2(\Omega)$ and $u(t_0) = 0$.

The proof of Theorem 1 rests upon the following theorem on non-linear functional equations in reflexive Banach spaces.

THEOREM 2. *Let X be a reflexive separable Banach space, X^* its conjugate space, (w,u) the pairing between elements w of X^* and u of X. Let T be an operator, not necessarily linear, with dense domain $D(T)$ in X and range $R(T)$ in X^*. Suppose that $T=L+G$, where:*

a) *L is a closed densely defined linear operator from X to X^* such that if L^* is its adjoint operator (also acting from a dense domain in X to X^*) and if $W=D(L)\cap D(L^*)$, then L is the closure of its restriction to W and L^* is the closure of its restriction to W.*

b) *G is an operator from X to X^* defined on all of X, maps bounded sets of X into bounded sets of X^*, and is demi-continuous, (i.e. continuous from the strong topology of X to the weak topology of X^*).*

c) *There exists a completely continuous mapping C of X into X^* such that for all u, v of $D(T)$*

$$\text{(1.11)} \qquad \operatorname{Re}(Tu-Tv,u-v)\geqq-\operatorname{Re}(Cu-Cv,u-v)$$

d) *There exists a continuous real-valued function $c(r)$ of the real variable r with $c(r)\to+\infty$ as $r\to+\infty$, such that for all u in $D(T)$,*

$$\text{(1.12)} \qquad \operatorname{Re}(Tu,u)\geqq c(\|u\|)\|u\|.$$

Then T is onto X^*, i.e. for every w of X^* there exists u in $D(T)$ such that $Tu=w$.

The proof of Theorem 2 is given in Section 2, that of Theorem 1 in Section 3. We observe that our present Theorem 2 includes as a special case Theorem 2 of [6].

Section 2. We now proceed to the proof of Theorem 2 which is based on the series of Lemmas which follows:

Definition 2.1. *Let T be a mapping with domain a linear subset $D(T)\subset X$, and range in X^*. Then T is said to be hemi-continuous if T is continuous from every closed line segment in $D(T)$ to the weak topology in X^*.*

LEMMA 2.1. *Let T be a hemi-continuous mapping from the dense linear subset $D(T)$ of X to X^*. Suppose that for some u_0 in $D(T)$ and w in X^*, we have*

(2.1) $$\operatorname{Re}(w - Tu, u_0 - u) \geqq 0$$

for all u in $D(T)$.

Then $w = Tu_0$.

Proof of Lemma 2.1. Suppose $w - Tu_0 \neq 0$. Then there exists v in the dense subset $D(T)$ such that

$$\operatorname{Re}(w - Tu_0, v) > 0.$$

Let $u_t = u_0 + tv$ for $t > 0$. Then $u_t \in D(T)$ for all t and replacing u in inequality (2.1) by u_t, we have $0 \leqq \operatorname{Re}(Tu_t - w, tv) = t \operatorname{Re}(Tu_t - w, v)$.

Cancelling t, we have

$$\operatorname{Re}(Tu_t - w, v) \geqq 0.$$

We rewrite this as

$$\operatorname{Re}(Tu_t - Tu_0, v) \geqq \operatorname{Re}(w - Tu_0, v).$$

Letting $t \to 0$, the lefthand side goes to zero by the hemi-continuity of T. Hence

$$0 \geqq \operatorname{Re}(w - Tu_0, v) > 0,$$

which is a contradiction. Q. E. D.

LEMMA 2.2. *Under the hypotheses of Theorem 2, $R(T)$ is closed in X^*.*

Proof of Lemma 2.2. Let $\{u_k\}$ be an infinite subsequence from X such that $Tu_k \to w$ strongly in X^*. Then $\| Tu_k \| \leqq M_0$ for all k. Further,

$$\operatorname{Re}(Tu_k, u_k) \geqq c(\| u_k \|) \| u_k \|,$$

while

$$\operatorname{Re}(Tu_k, u_k) \leqq \| Tu_k \| \cdot \| u_k \| \leqq M_0 \| u_k \|.$$

Thus

$$c(\| u_k \|) \leqq M_0$$

and since $c(r) \to \infty$ as $r \to +\infty$, $\| u_k \| \leqq M_1$. Since G carries bounded sets in X into bounded sets in X^*, we know that $\| Gu_k \| \leqq M_2$. Replacing $\{u_k\}$ by an infinite subsequence if necessary, we may assume that $u_k \to u_0$ weakly in X, $Gu_k \to v_0$ weakly in X^*. Hence $Lu_k = Tu_k - Gu_k \to w - v_0$ weakly in X^*.

The graph of L is closed in $X \times X^*$ since L is a closed linear operator. Being a closed linear subset of $X \times X^*$, the graph of L is weakly closed. Hence the ordered pair $u_0, w - v$ lies in the graph of L, i. e. $u_0 \in D(L)$.

Since C is compact, $Cu_k \to Cu_0$ strongly in X^*. For each value of k and each v in $D(T)$

$$\operatorname{Re}(Tu_k - Tv, u_k - v) \geqq -\operatorname{Re}(Cu_k - Cv, u_k - v).$$

Since $Tu_k - Tv$ converges strongly to $w - Tu$ while $u_k - v$ converges weakly to $u_0 - v$,

$$\operatorname{Re}(Tu_k - Tv, u_k - v) \to \operatorname{Re}(w - Tu, u_0 - v).$$

Similarly

$$\operatorname{Re}(Cu_k - Cv, u_k - v) \to \operatorname{Re}(Cu_0 - Cv, u_0 - v).$$

Hence if $T_1 = T + C$,

$$\operatorname{Re}(w + Cu_0 - T_1 v, u_0 - v) \geqq 0$$

for all $v \in D(T) = D(T_1)$. Since T_1 is obviously hemi-continuous since T is hemi-continuous and since $u_0 \in D(L) = D(T_1)$, it follows from Lemma (2.1) that

$$w + Cu_0 = T_1 u_0 = Tu_0 + Cu_0$$

i. e. $w = Tu_0$. Q. E. D.

Lemma 2.3. *Let P be a projection of X onto a finite dimensional subspace F with $F \subset D(T)$, P^* the adjoint projection of X^*, $F' =$ the range of P^*. Then if T is continuous from finite dimensional subspaces of X to the weak topology of X^* and if*

$$\operatorname{Re}(Tu, v) \geqq c(\| u \|) \| u \|$$

for all u in $D(T)$ where $c(r) \to +\infty$ as $r \to +\infty$, then the mapping P^*T maps F onto F'.

Proof of Lemma 2.3. Lemma 2.3 is identical with Lemma 4 of [6].

Lemma 2.4. *Let W be a dense linear subset of X and consider the family $\{P_\eta\}$ of all projections of X onto finite dimensional subspaces contained in W. Let $F' = \bigcup_\eta R(P_\eta{}^*)$. Then F' is dense in X^*.*

Proof of Lemma (2.4). If F' is not dense in X^*, then by the reflexivity of X there exists u_0 in X, $u_0 \neq 0$, which annihilates F', i. e.

$$(P_\eta{}^* v, u_0) = 0 \tag{4.2}$$

for all v in X^* and all η. Equation (4.2) is equivalent to $(v, P_\eta u_0) = 0$, i. e. $P_\eta u_0 = 0$ for all η.

Since W is dense in X, for each $\epsilon > 0$ there exists u_1 in W such that $\| u_0 - u_1 \| < \epsilon$. Since both u_0 and u_1 are non-null if $\epsilon < \| u_0 \|$, by the Hahn-Banach theorem we may choose v in X^* with $\| v \| = \dfrac{1}{\| u_1 \|}$ and $(v, u_1) = 1$.

Let P be the projection of X onto the one-dimensional space spanned by u_1 given by

$$Pu = (v, u)u_1$$

for $u \in X$. Then

$$\begin{aligned} \| Pu_0 \| &= \| u_1 + (v, u_0 - u_1)u_1 \| \\ &\geqq \| u_1 \| - \| v \| \cdot \| u_0 - u_1 \| \cdot \| u_1 \| = \| u_1 \| - \epsilon. \end{aligned}$$

Choosing $\epsilon < \| u_0 \|/2$, we have

$$\| u_1 \| \geqq \| u_0 \| - \| u_1 - u_0 \| \geqq \| u_0 \| - \frac{\| u_0 \|}{2} > \epsilon.$$

For such ϵ, $\| Pu_0 \| > 0$ and $Pu_0 \neq 0$ contradicting the fact that all finite dimensional projections annihilate u_0. Q. E. D.

Proof of Theorem 2. By Lemma 2. 2, $R(T)$ is closed in X^*. Let W_1 be a dense subset of X, and let $F' = \bigcup_{\eta} R(P_\eta{}^*)$ where P_η ranges over the projections of X on finite dimensional subspaces of W_1. By Lemma 2. 4, each such F' is dense. Hence it suffices to show that for some suitable W_1, $F' \subset R(T)$.

We may choose a denumerable subset W_1 of $D(L) \cap D(L^*)$ such that L is the closure of L restricted to W_1 and L^* is the closure of L^* restricted to W. Let $F' = \bigcup_{\eta} R(P_\eta{}^*)$ where P_η ranges over the projections of X on finite dimensional subspaces of W_1. Let $w \in F'$. Then there exists a projection which we shall call P_1 of X on a finite-dimensional subspace F_1 of W_1 such that $P_1{}^*w = w$. We construct an increasing family of finite-dimensional subspaces F_j of W_1 starting with F_1 such that

$$W_1 = \bigcup_{j} F_j.$$

By Lemma 3 of [7], we may find a sequence of projections P_j of X on F_j such that

$$P_jP_k = P_kP_j = P_j, \quad j < k.$$

Moreover, if $\{P_j{}^*\}$ is the sequence of adjoint projections of X^*, and if $F'_j = R(P_j{}^*)$, then $F'_j \subset F'_{j+1}$ and in particular $w \in F'_1 \subset F'_j$ for every $j \geqq 1$.

Since $W_1 \subset D(L) = D(T)$, the mapping $P_j{}^*T$ maps each F_j into F'_j and indeed onto F'_j by Lemma 2. 3. Hence we may choose $u_j \in F_j$ such that

$$P_j{}^*Tu_j = w.$$

For such u_j, we know that

$$\operatorname{Re}(w,u_j)=\operatorname{Re}(P_j{}^*Tu_j,u_j)=\operatorname{Re}(Tu_j,u_j)\geqq c(\|u_j\|)\|u_j\|.$$

and hence

$$c(\|u_j\|)\leqq\|w\|.$$

Since $c(r)\to\infty$ as $r\to\infty$, we know that $\|u_j\|\leqq M$ for all j and some constant M. Hence $\|Gu_j\|\leqq M_1$ for all j. Passing to a subsequence if necessary, we may assume that $u_j\to u_0$ weakly in X, $Gu_j\to v_0$ weakly in X^*.

Let v be any element of W_1. Then v lies in F_r for some r, so that $P_jv=v$ for $j\geqq r$.

For $j\geqq r$

$$(w,v)=(P_j{}^*Tu_j,v)=(Tu_j,v)+(Lu_j,v)+(Gu_j,v).$$

Since $v\in W_1\subset D(L^*)$,

$$(Lu_j,v)=(u_j,L^*v)\to(u_0,L^*v).$$

Since $Gu_j\to v_0$ weakly in X^*,

$$(Gu_j,v)\to(v_0,u).$$

Thus

$$(w,v)=(u_0,L^*v)+(v_0,v)$$

i. e.

$$(u_0,L^*v)=(w-v_0,v)$$

for all $v\in W_1$. However, both sides of the last equation are continuous in the graph of L^* and the graph of L^* restricted to W_1 is dense in the graph of L^* by hypothesis. Hence

$$(u_0,L^*v)=(w-v_0,v)$$

for all $v\in D(L^*)$, i. e. $u_0\in D(L)$ and $Lu_0=w-v_0$.

Finally, let $v\in W_1$. Then $v\in F_r$ for some r. For $j\geqq r$,

$$\begin{aligned}(Tu_j,u_j-v)&=(Tu_j,P_j(u_j-v))=(P_j{}^*Tu_j,u_j-v)\\&=(w,u_j-v)\to(w,u_0-v)\end{aligned}$$

as $j\to+\infty$. For all j, we know that

$$\begin{aligned}-\operatorname{Re}(Cu_j-Cv,u_j-v)&\leqq\operatorname{Re}(Tu_j-Tv,u_j-v)\\&=\operatorname{Re}(Tu_j,u_j-v)-\operatorname{Re}(Tv,u_j-v)\end{aligned}$$

where

$$\operatorname{Re}(Cu_j-Cv,u_j-v)\to\operatorname{Re}(Cu_0-Cv,u_0-v)$$

since $u_j \to u_0$ weakly in X, and $Cu_j \to Cu_0$ strongly in X^* by the complete continuity of C. In addition

$$\operatorname{Re}(Tu_j, u_j - v) - \operatorname{Re}(Tv, u_j - v) \to \operatorname{Re}(w, u_0 - v) - \operatorname{Re}(Tv, u_0 - v).$$

Hence for $v \in W_1$

$$-\operatorname{Re}(Cu_0 - Cv, u_0 - v) \leqq \operatorname{Re}(w - Tv, u_0 - v).$$

Setting $T_1 = T + C$, this last inequality may be written as

$$\operatorname{Re}(w + Cu_0 - T_1 v, u_0 - v) \geqq 0$$

for all v in W_1.

Let v be any element of $D(T_1) = D(L)$. Then there exists a sequence $\{v_k\}$ from W_1 such that $v_k \to v$ strongly in X, $Lv_k \to Lv$ strongly in X^*. Since G is demi-continuous, $Gv_k \to Gv$ weakly in X^*. Since C is compact, and hence continuous, $Cu_k \to Cv$ strongly in X^*. Hence $T_1 v_k \to T_1 v$ weakly in X^*. For each k, we have

$$\operatorname{Re}(w + Cu_0 - T_1 v_k, u_0 - v_k) \geqq 0.$$

Since $u_0 - v_k \to u_0 - v$ strongly in X, $w + Cu_0 - T_1 v_k \to w + Cu_0 - T_1 v$ weakly in X^*, we have in the limit

$$\operatorname{Re}(w + Cu_0 - T_1 v, u_0 - v) \geqq 0$$

for all v in $D(T_1)$. Applying Lemma (2.1) once more, we see that $T_1 u_0 = w + Cu_0$ i.e. $Tu_0 = w$. Hence

$$F' \subset R(T). \qquad \text{Q.E.D.}$$

Section 3. We now pass to the proof of Theorem 1. As a technical device in this proof we may define a global boundary value problem on the infinite cylinder $\Omega \times R^1$ as follows:

Let $L^p(R^1, V)$ be the Banach space formed by the p-th-power summable functions from R' to V with norm

$$\| u \| = \{ \int_{-\infty}^{+\infty} \| u(t) \|^p_{m,p} \}^{1/p}.$$

$F^* = \{ u \mid u \in L^p(R', V)$, u is continuous from R^1 to V, u is C^1 from R^1 to $L^q(\Omega)$, u has compact support in $R^1\}$.

For each u in F^*, consider the sesqui-linear form

$$k(u, v) = \int_{-\infty}^{+\infty} \langle \frac{du}{dt}(t), v(t) \rangle \, dt$$

defined for all $v \in L^p(R^1, V)$. Then $k(u,v)$ is a bounded conjugate-linear functional of v for fixed u. Hence we may define:

Definition. For each u in $F^\#$ $L_0 u$ is the element of $(L^p(R^1,V))^*$ such that

$$(3.2) \qquad (L_0 u, v) = k(u,v) = \int_{-\infty}^{+\infty} \langle \frac{du}{dt}(t), v(t) \rangle \, dt,$$

for all v in $L^p(R^1, V)$.

By an integration by parts, we see that $-L_0 \subseteq L_0^*$. Hence L_0 is closable as an operator from $L^p(R^1,V)$ to $(L^p(R^1,V))^*$.

Definition 3.1. *L is the closure of L_0 as an operator with domain in $L^p(S,V)$ and range in $(L^p(S,V))^*$.*

Global variational boundary value problem: Let $f \in (L^p(R^1,V))^*$. Then $u \in L^p(R^1,V)$ is said to be a solution of the global variational boundary value problem for the equation

$$\frac{\partial u}{\partial t} + A(t)u = f$$

on $\Omega \times R^1$ if $u \in D(L)$ and

$$(3.3) \qquad (Lu, v) + h(u,v) = (f, v)$$

for all v in $L^p(R^1,V)$, where

$$(3.4) \qquad h(u,v) = \int_{-\infty}^{+\infty} \sum_{|\alpha| \leq m} \langle A_\alpha(x,t,u,\cdots,D^m u), D^\alpha v(t) \rangle \, dt.$$

LEMMA 3.1. $L^* = -L$.

Proof of Lemma 3.1. We know that $-L \subseteq L^*$ since $-L_0 \subseteq L^*$ and L^* is a closed operator. To show the reverse, suppose $u \in D(L^*)$, $L^* u = u_1$. Then for all v in $D(L_0)$

$$(u, L_0 v) = (u_1, v)$$

i. e.

$$(3.5) \qquad \int_{-\infty}^{+\infty} \langle u(t), \frac{dv}{dt}(t) \rangle \, dt = \int_{-\infty}^{+\infty} \langle u_1(t), v(t) \rangle \, dt.$$

Let $j \in C_0^\infty(R^1)$ with $\int_{-\infty}^{+\infty} j(s)\, ds = 1$, and set

$$j_\epsilon(s) = \epsilon^{-1} j(s/\epsilon), \quad \epsilon > 0$$

$$J_\epsilon u(t) = \int_{-\infty}^{+\infty} j_\epsilon(t-s) u(s)\, ds.$$

Using (3.5), we find using Fubini's theorem that

$$\int_{-\infty}^{+\infty} \langle (J_\epsilon u)(t), \frac{dv}{dt}(t)\rangle \, dt = \int_{-\infty}^{+\infty} \langle (J_\epsilon u_1)(t), v(t)\rangle \, dt \tag{3.6}$$

i. e.

$$(J_\epsilon u, L_0 v) = (J_\epsilon u_1, v).$$

Hence $L^* J_\epsilon u = J_\epsilon u_1$. Since $J_\epsilon u \to u$ in $L^p(R^1, V)$, $J_\epsilon u_1 \to u_1$ in $L^p(R^1, V)$, it suffices to show that $J_\epsilon u \in D(L)$. The latter fact follows if we choose ϕ in $C_0^\infty(R^1)$ with $\phi(s) = 1$ for $|s| \leqq 1$, set $\phi_\epsilon(s) = \phi(\epsilon s)$, and let

$$v_\epsilon(s) = \phi_\epsilon(s) J_\epsilon u(s).$$

Then as $\epsilon \to 0$, $v_\epsilon \to J_\epsilon u$ in $L^2(R, V)$, $\frac{dv_\epsilon}{ds} \to \frac{d(J_\epsilon u)}{ds}$ in $L^2(R, V)$. Since $v_\epsilon \in D(L_0)$ for each $\epsilon > 0$, and $L_0^* J_\epsilon u = -\frac{d(J_\epsilon u)}{ds}$ by an obvious integration by parts, it follows that $J_\epsilon u \in D(L)$ and the proof of Lemma (3.2) is complete.

LEMMA 3.2. *If $u \in D(L)$ with $u(s) = 0$ for $s < t_0$ and if S is the interval $[t_0, t_1]$, δ the corresponding operator defined in Section 1, then the restriction u_0 of u to S lies in $D(\delta)$ and $\delta(u_0) = v_0$ where v_0 is the restriction of $v = Lu$ to S.*

Proof of Lemma 3.2. Let w be a continuously differentiable function from R^1 to V such that $w(s) = 0$ for $s \geqq t_1$ and $w(s) = 0$ for $s \leqq t_0 - \delta$ for some $\delta > 0$. Then w lies in $D(L_0)$ and hence

$$(u, L_0 w) = -(Lu, w).$$

i. e.

$$\int_{-\infty}^{+\infty} \langle u(t), \frac{dw}{dt}(t)\rangle \, dt = -\int_{-\infty}^{+\infty} \langle v(t), w(t)\rangle \, dt. \tag{3.7}$$

Since $u(t) = 0$ for $t < t_0$, while $\frac{du}{dt}(t) = 0$ for $t > t_1$, the integral on the left equals

$$\int_{t_0}^{t_1} \langle u(t), \frac{dw}{dt}(t)\rangle \, dt = \int_{t_0}^{t_1} \langle u_0(t), \frac{dw}{dt}(t)\rangle \, dt.$$

Since $u(t) = 0$ for $t < t_0$, and $v = \frac{du}{dt}$ in the distribution sense, $v(t) = 0$ for $t < t_0$. Hence the integral on the right side of equation (3.7) equals

$$-\int_{t_0}^{t_1}\langle v(t), w(t)\rangle\,dt = -\int_{t_0}^{t_1}\langle v_0(t), w(t)\rangle\,dt.$$

Hence

$$-\int_{t_0}^{t_1}\langle u_0(t), \frac{dw}{dt}\rangle\,dt = \int_{t_0}^{t_1}\langle v_0(t), w(t)\rangle\,dt.$$

In terms of the operator δ_1 of Section 1, we have

$$(u_0, \delta_1 w) = (v_0, w)$$

for all w in $D(\delta_1)$. Hence u_0 lies in $D(\delta_1{}^*)$ and $\delta_1{}^* u_0 = v_0$. It follows from essentially the same argument as used in the proof of Lemma (3.1), however, that $\delta_1{}^* = \delta$. Hence $u_0 \in D(\delta)$ and $\delta u_0 = v_0$. Q.E.D.

Lemma 3.3. *Suppose $u \in D(L)$. Then after a change on a set of measure zero on R^1, u is a continuous function from R^1 to $L^2(\Omega)$. For every s in R^1, $u(s)$ then satisfies the equation*

$$\|u(s)\|^2_{L^2} = 2\,\mathrm{Re}(Lu, T_s u) \tag{3.8}$$

where $T_s u = \mathcal{X}_s \cdot u$ and $\mathcal{X}_s$ is the characteristic function of the interval $(-\infty, s)$.

Proof of Lemma (3.3). For $u \in D(L_0)$, we have

$$\|u(s)\|^2_{L^2(\Omega)} = 2\,\mathrm{Re}\int_{-\infty}^{s}\langle \frac{du}{dt}(t), u(t)\rangle\,dt = 2\,\mathrm{Re}(L_0 u, T_s u) \tag{3.9}$$
$$\leqq 2\,\|L_0 u\|_{X^*}\,\|T_s u\|_X,$$

where $X = L^p(R^1, V)$. Since L is the closure of L_0 as a mapping from X to X^*, for each u in $D(L)$ there exists a sequence $\{u_k\}$ from $D(L_0)$ such that $u_k \to u$ in X while $L_0 u_k \to Lu$ in X^*. For each k and r, we have

$$\|u_k(s) - u_r(s)\|^2_{L^2(\Omega)} \leqq 2\,\|L_0(u_k - u_r)\|_{X^*}\,\|T_s(u_k - u_r)\|_X$$
$$\leqq 2\,\|L_0 u_k - L_0 u_r\|_{X^*}\,\|u_k - u_r\|_X \to 0$$

as $k, r \to \infty$. Hence $u_k(s)$ converges in $L^2(\Omega)$ uniformly in s on R^1 to a uniformly continuous function from R^1 to $L^2(\Omega)$. Since, $u_k \to u$ in $L^p(R^1, V)$, it follows that u is uniformly continuous from R^1 to $L^2(\Omega)$. Moreover, equation (3.3) follows from equation (3.9) by continuity since $\|u_k(s)\|_{L^2(\Omega)} \to \|u(s)\|_{L^2(\Omega)}$ and $(L_0 u_k, T_s u_k) \to (Lu, T_s u)$ as $k \to \infty$. Q.E.D.

Lemma 3.4. *If $f \in (L^p(R^1, V))^* = L^q(R^1, V^*)$ with $f(t) = 0$ a.e. for $t < t_0$ and if u is a solution of the global boundary problem for $\frac{}{\partial t} + A(t)u = f$ with respect to V, then $u(t) = 0$ for $t < t_0$.*

Proof of Lemma 3.4. By Lemma (3.3), for $t < t_0$,

$$\| u(t) \|_{L^2} = 2 \operatorname{Re}(Lu, T_t u)$$

where $T_t u = \mathcal{X}_t \cdot u$. If u is a solution of the global boundary problem, then for every v in $L^p(R^1, V)$, we have

$$(Lu, v) + h(u, v) = (f, v).$$

Let $v = T_t u$. Then

$$(f, v) = (f, T_t u) = 0$$

since $f(s) = 0$ for $s < t_0$ and t is less than t_0. Furthermore

$$\begin{aligned} h(u, T_t u) &= \int_{-\infty}^{+\infty} a(s; u(s), \mathcal{X}_t(s) u(s))\, ds \\ &= \int_{-\infty}^{t} a(s, (\mathcal{X}_t u)(s), (\mathcal{X}_t u)(s))\, ds \\ &= h(T_t u, T_t u). \end{aligned}$$

Hence $\operatorname{Re} h(T_t u, T_t u) \geqq 0$.

Finally

$$\operatorname{Re}(Lu, T_t u) = -\operatorname{Re}\{h(u, T_t u) + (f, T_t u\} \leqq 0.$$

Thus $\| u(t) \|^2 \leqq 0$, which implies that $u(t) = 0$. Q.E.D.

PROPOSITION 3.1. *Let* $f_0 \in (L^p(S, V))^*$ *and let* $f \in (L^p(R^1, V))^*$ *be obtained by setting* $f = f_0$ *on* S, $f = 0$ *outside* S. *If* u *is a solution of the global problem for* $\frac{\partial u}{\partial t} + A(t) u = f$ *with respect to* V, *then the restriction* u_0 *of* u *to* S *is a solution of the variational boundary value problem for*

$$\frac{\partial u}{\partial t} + A(t) u = f_0$$

with respect to V *on* $\Omega \times S$.

If global solutions exists for all f, *then solutions* u_0 *will exist for all* f_0.

Proof of Proposition 3.1. This follows directly from the preceding lemmas.

It follows from Assumption (I) on $A(t)$

$$| h(u, v) | \leqq c \| v \| \{ \| u \| + 1 \}^{p-1} \tag{3.10}$$

for all u, v in $L^p(R^1, V)$. Thus for fixed u in $L^p(R^1, V)$, $h(u, v)$ is a bounded conjugate-linear functional of v in $L^p(R^1, V)$.

Definition 3.2. *For each u in $L^p(R^1, V)$ Gu is the unique element of $(L^p(R^1, V))^*$ such that*

$$(Gu, v) = h(u, v)$$

for all v in $L^p(R^1, V)$.

From (3.10), we see that

$$\| Gu \| \leqq C\{\| u \| + 1\}^{p-1}.$$

so that G carries bounded sets of $L^p(R^1, V)$ into bounded sets of $(L^p(R^1, V))^*$.

LEMMA 3.5. *u is a solution of the global boundary value problem with a homogeneous term f if and only if $u \in D(L)$ and*

$$(L + G)u = f.$$

Proof of Lemma 3.5. This is simply a translation of the definition of the global boundary value problem using the definition of L and G.

LEMMA 3.6. *There is at most one solution u of the variational boundary value problem on the interval S for the equation $\frac{\partial u}{\partial t} + A(t)u = f$ with respect to V for a given f in $(L^p(S, V))^*$.*

Proof of Lemma (3.6). As in the proof of Lemma (3.4), we may show that every u in $D(\delta)$ may be chosen continuous and for every $t_0 \leqq t \leqq t_1$,

$$\| u(t) \|^2_{L^2} = 2 \operatorname{Re}(\delta u, T_t u) \tag{3.11}$$

where $T_t u = \mathcal{X}_t \cdot u$ and $\mathcal{X}_t$ is the characteristic function of the interval $[t_0, t]$. Let u and v be two solutions of the variational boundary problem for a given f. Then we have for all w in $L^2(S, V)$

$$(\delta u, w) + h(u, w) = (f, w)$$
$$(\delta v, w) + h(v, w) = (f, w)$$

Subtracting, we have

$$(\delta(u - v), w) = - \{h(u, w) - h(v, w)\}.$$

Applying equation (3.11) with u replaced by $(u - v)$, we have

$$\begin{aligned} \| u(t) - v(t) \|^2_{L^2} &= -2 \operatorname{Re}(\delta(u - v), T_t(u - v)) \\ &= -2 \operatorname{Re}\{h(u, T_t(u - v)) = h(v, T_t(u - v))\}. \end{aligned}$$

However, $h(u, T_t(u - v)) = h(T_t u, T_t(u - T_t v)$ while

$$h(v, T_t(u - v)) = h(T_t v, T_t u - T_t v).$$

Hence,

$$\| u(t) - v(t) \|^2_{L^2}$$
$$= -2\,\mathrm{Re}\{h(T_t u, T_t u - T_t v) - h(T_t v, T_t u - T_t v)\} \leqq 0$$

so that $\| u(t) - v(t) \|_{L^2} = 0$ for $t \leqq t_1$, and finally $u = v$. Q.E.D.

LEMMA 3.7. *G is demi-continuous, i.e. it is continuous from the strong topology of $L^p(R^1, V)$ to the weak topology of $(L^p(R^1, V))^*$.*

Proof of Lemma 3.7. We have already observed that G maps bounded sets of $L^p(R^1, V)$ into bounded sets of $(L^p(R^1, V))^*$. Let $\{u_k\}$ be a sequence from $L^p(R^1, V)$. If $\{Gu_k\}$ does not converge weakly to Gu in $(L^p(R^1, V))^*$, there exists an infinite subsequence which we may again designate as $\{Gu_k\}$ and an element v of $L^p(R^1, V)$ such that (Gu_k, v) converges to a limit different from (Gu, v). Hence it suffices to show that this last subsequence (Gu_k, v) has an infinite subsequence converging to (Gu, v).

Since $u_k \to u$ strongly in $L^p(S, V)$, we know that for $k \to \infty$

$$\int_{-\infty}^{+\infty} \int | D^\alpha u_k - D^\alpha u |^p \, dx dt \to 0.$$

We therefore choose an infinite subsequence, which we again denote by $\{u_k\}$, such that $D^\alpha u_k \to D^\alpha u$ a.e. on $\Omega \times R^1$.

Let

$$h_k = \sum_{|\beta| \leqq m} | D^\beta u_k |^{p-1} + g(x, t), \tag{3.12}$$

$$h = \sum_{|\beta| \leqq m} | D^\beta u |^{p-1} + g(x, t), \tag{3.13}$$

$$g_{\alpha,k} = \frac{1}{h_k} A_\alpha(x, t, u_k, \cdots, D^m u_k), \tag{3.14}$$

$$g_\alpha = \frac{1}{h} A_\alpha(x, t, u, \cdots, D^m u). \tag{3.15}$$

We may obviously assume without loss of generality that $g(x, t) > 0$ for all (x, t) in $\Omega \times R^1$. Thus all the above are well-defined and we know from Assumption (I) and the a.e. convergence of $D^\beta u_k$ to $D^\beta u$ for $| \beta | \leqq m$, that

$$| g_{\alpha,k} | \leqq M \text{ for all } k \text{ and } \alpha \text{ with } | \alpha | \leqq m,$$
$$g_{\alpha,k} \to g_\alpha \text{ a.e. for } | \alpha | \leqq m.$$

Moreover for all k, h_k and h lie in $L^q(\Omega \times R^1)$, $q = p(p-1)^{-1}$. Furthermore

$$\int_{S_1} \int_\Omega | (h_k - h)(x, t) |^q \, dx dt \to 0$$

as $k \to \infty$.

Thus

$$(Gu_k, v) - G(u, v) = \sum_\alpha \int_{-\infty}^{+\infty} \langle A_\alpha(x, t, u_k, \cdots, D^m u_k) \\ - A_\alpha(x, t, u, \cdots, D^m u), D^\alpha v\rangle\, dt$$

$$= \sum_\alpha \{ \int_{-\infty}^{+\infty} \langle g_{\alpha k} h_k, D^\alpha v\rangle\, dt - \int_{-\infty}^{+\infty} \langle g_\alpha h, D^\alpha v\rangle\, dt \}$$

However

$$\int_{-\infty}^{+\infty} \langle g_{\alpha,k} h_k, D^\alpha v\rangle\, dt = \int_\Omega \int_{R_1} h_k(x, t) \{ g_{\alpha,k}(x, t) D^\alpha v(x, t)\}^* \, dx dt$$

where $h_k \to h$ strongly in $L^q(\Omega \times R^1)$ while $g_{\alpha,k} D^\alpha v$ converges strongly to $g_\alpha D^\alpha v$ strongly in $L^p(\Omega \times R^1)$ by dominated convergence. Hence

$$\int_{-\infty}^{+\infty} \langle g_{\alpha,k} h_k, D^\alpha v\rangle\, dt \to \int_{-\infty}^{+\infty} \langle g_\alpha h, D^\alpha v\rangle\, dt$$

and $(Gu_k, v) \to G(u, v)$. Q.E.D.

Proof of Theorem 1. We know from Lemma 3.6 that there can be at most one solution of our given boundary value problem. We only need to establish the existence of at least one solution u. By Proposition (3.1), we need only show that the global boundary problem has a solution for every f in $(L^p(R^1, V))^*$. By Lemma 3.5, this will be true if and only if for the operators L and G defined previously in this Section, $T = L + G$ maps $L^p(R^1, V)$ *onto* $(L^p(R^1, V))^*$.

By Lemma 3.7, G is demi-continuous and maps bounded sets of $L^p(R^1, V)$ onto bounded sets of $(L^p(R^1, V))^*$. L is a closed linear operator with dense domain in $L^p(R^1, V)$. By Lemma 3.1, $L^* = -L$ so that $D(L) \cap D(L^*) = D(L)$ and the condition that L is the closure of its restriction to $D(L) \cap D(L^*)$ is satisfied vacuously. Finally,

$$\operatorname{Re}(Tu - Tv, u - v) = \operatorname{Re}(L(u - v), u - v) + \operatorname{Re}(Gu - Gv, u - v)$$

where

$$\operatorname{Re}(L(u - v), u - v) = -\operatorname{Re}(u - v, L(u - v)) = -\operatorname{Re}(L(u - v), u - v)$$

so that

$$\operatorname{Re}(L(u - v), u - v) = 0,$$

while

$$\operatorname{Re}\{ (Gu, u - v) - (Gu, u - v)\} = \operatorname{Re}\{ h(u, u - v) - h(v, u - v)\} \geqq 0$$

for all $u, v \in L^p(R^1, V)$. Hence

$$\operatorname{Re}(Tu - Tv, u - v) \geqq 0$$

for all $u, v \in D(T)$.

Similarly for $u \in D(T)$

$$\operatorname{Re}(Tu, v) = \operatorname{Re}(Gu, v) - \operatorname{Re} h(u, v) \geqq c(\| u \|) \| u \|$$

so that all the hypotheses of Theorem 2 are satisfied for T. Hence T is onto. Q. E. D.

INSTITUTE FOR ADVANCED STUDY
AND
UNIVERSITY OF CHICAGO.

REFERENCES.

[1] F. E. Browder, "Functional analysis and partial differential equations, I," *Mathematische Annalen*, vol. 138 (1959), pp. 55-79.

[2] ———, "On the solvability of non linear functional equations," *Duke Mathematical Journal*, vol. 30 (1963).

[3] ———, "Variational boundary value problems for quasi-linear elliptic equations of arbitrary order," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 50 (1963), pp. 31-37.

[4] ———, "Variational boundary value problems for quasi-linear elliptic equations, II," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 50 (1963), pp. 592-598.

[5] ———, "Variational boundary value problems for quasi-linear elliptic equations, III," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 50 (1963), pp. 794-798.

[6] ———, "Non-linear elliptic boundary value problems," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 862-874.

[7] ———, "Non-linear equations of evolution," to appear in *Annals of Mathematics*.

[8] ———, "Non-linear parabolic boundary value problems of arbitrary order," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 858-861.

[9] J. L. Lions, *Equations differentielles operationnelles et problemes aux limites.* (Die Grundlehren der Mathematischen Wissenschaften, Bd. 111) Berlin-Göttingen-Heidelberg, 1961.

[10] G. J. Minty, "Monotone (non-linear) operators in Hilbert space," *Duke Mathematical Journal*, vol. 29 (1962), pp. 341-346.

[11] M. M. Vainberg and R. I. Kachurovski, "On the variational theory of non-linear operators and equations," *Doklady Akademii Nauk SSSR*, vol. 129 (1959), pp. 1199-1202.

[12] M. I. Visik, "Solution of a system of quasi-linear equations in divergence form under periodic boundary conditions," *Doklady Akademii Nauk SSSR*, vol. 137 (1961), pp. 502-505.

[13] ———, "Boundary value problems for quasi-linear strongly elliptic systems in divergence form," *Doklady Akademii Nauk SSSR*, vol. 138 (1961), pp. 518-521.

[14] ———, "Simultaneous quasi-linear elliptic equations with lower order terms," *Doklady Akademii Nauk SSSR*, vol. 144 (1962), pp. 13-16.

COMPACTIFICATIONS OF SYMMETRIC SPACES II: THE CARTAN DOMAINS.*

By CALVIN C. MOORE.

Section 1. If G is a semi-simple Lie group with finite center and K a maximal compact subgroup, the space G/K is a symmetric Riemannian space of non-compact type. In a previous paper [13], we have investigated certain compactifications of G/K which are obtained by embedding G/K into a space of probability measures on a boundary of G. We have shown that these compactifications, introduced by Furstenberg [2], are the same as Satake's compactifications [15]. Satake remarks (p. 104) that if G/K is a classical Cartan domain [1], then one can verify case by case that one of these compactifications is homeomorphic to the closure of the Cartan domain when it is viewed as a bounded domain in complex Euclidean space. One of our main results will be to give an intrinsic proof of this theorem. In the course of the proof we will be able to construct the desired homeomorphism in a very natural way using the properties of Furstenberg's compactifications.

Section 2 is devoted to the proof of an algebraic result concerning the roots of a Cartan domain which is needed later. This theorem has been conjectured by Bott and Koranyi. Although one could verify this result also by checking it case by case, it seems to be of interest to have an intrinsic proof. In Section 3 we prove the theorem mentioned above. Lemma 3.1 on the structure of the Shilov boundary of the domain has been proved by Koranyi and Wolf [11]. In Section 4 we shall discuss the boundary components of a Cartan domain as defined by Pyatetskiĭ-Shapiro [14]. We will give intrinsic proof of his main theorems and some additional information which we can obtain from our point of view.

Section 2. Let G be the connected component of the group of analytic automorphisms of an irreducible bounded symmetric domain D. Then $D \cong G/K$ where K is a maximal compact subgroup of G. For the following see for instance, [4], [5]. It is known that K has one dimensional center

---

Received September 10, 1963.

* Research supported in part by NSF grant G-18974.

and has the same rank as G. Let $\mathfrak{g}_0$ and $\mathfrak{k}_0$ be the Lie algebras of G and K and let $\mathfrak{p}_0$ be the orthogonal complement of $\mathfrak{k}_0$ in $\mathfrak{g}_0$; further let $\mathfrak{a}_0$ be a Cartan subalgebra of $\mathfrak{k}_0$. Let $\mathfrak{g}$, $\mathfrak{k}$, $\mathfrak{p}$ and $\mathfrak{a}$ be their complexifications and $\mathfrak{u}=\mathfrak{k}_0+i\mathfrak{p}_0$ a compact form of g. Now $\mathfrak{a}$ is a Cartan subalgebra of $\mathfrak{g}$ and the roots of $\mathfrak{g}$ fall into two exhaustive classes, $C=\{\alpha\colon X_\alpha\in\mathfrak{k}\}$ and $P=\{\alpha\colon X_\alpha\in\mathfrak{p}\}$, the compact and non-compact roots respectively where X_α are root vectors. The adjoint representation of $\mathfrak{k}$ on $\mathfrak{p}$ decomposes $\mathfrak{p}$ into two irreducible subspaces $\mathfrak{p}^\pm$, and $\mathfrak{p}^\pm$ is spanned by $\{X_\alpha\colon\alpha\in P^\pm\}$ where $P^\pm$ denote the positive (negative) roots in P in an appropriate oredering. If $\mu_1,\cdots,\mu_n$ are the fundamental roots of $\mathfrak{g}$ (with respect to an ordering fixed once and for all) we may assume that μ_1 is the unique non-compact fundamental root. The compact roots are those of the form $\sum n_i\mu_i$ $(i\geqq 2)$.

Lemma 2.1. *If* $\beta\in P^+$, *then* $\beta=\mu_1+\sum n_i\mu_i$ $(i\geqq 2)$.

Proof. Since X_{μ_1} is in $\mathfrak{p}^+$ and $\operatorname{ad}(\mathfrak{k})$ is irreducible on $\mathfrak{p}^+$, every X_β can be written as $c\cdot[X_{\alpha_1},[X_{\alpha_2},\cdots,[X_{\alpha_n},X_{\mu_1}]]\cdots]$ where c is a constant and $\alpha_1,\cdots,\alpha_n$ are compact roots. It follows that $\beta=\mu_1+\sum n_i\mu_i$ $(i\geqq 2)$ as desired.

Harish-Chandra [4] has constructed a family of strongly orthogonal roots $\gamma_1,\cdots,\gamma_m$ which play a fundamental role in our discussion. (Two roots, α and β, are strongly orthogonal when $\alpha\pm\beta$ are not roots.) The root γ_1 is the lowest root in P^+; that is μ_1, and γ_r is inductively the smallest root of P^+ strongly orthogonal to $\gamma_1,\cdots,\gamma_{r-1}$. Let V be the subspace of $\mathfrak{a}$ spanned by the γ's and π the orthogonal projection of $\mathfrak{a}$ onto V. Harish-Chandra has proved the following facts.

Lemma 2.2. *If* β *is a positive compact root then there are three possibilities:*

1.) $\pi(\beta)=0$ *in which case* β *is strongly orthogonal to all* γ_i.

2.) $\pi(\beta)=-\frac{1}{2}\gamma_i$ *in which case* β *is strongly orthogonal to all* γ_k, $k\neq i$.

3.) $\pi(\beta)=\frac{1}{2}(\gamma_j-\gamma_i)$, $j>i$ *in which case* β *is strongly orthogonal to all* γ_k, $k\neq i,j$.

Moreover $\beta+\gamma_j$ *is not a root.*

We shall examine which of these combinations occur; the end results are Theorem 1 and 2. We first need some preliminary facts (D will denote the set of compact fundamental roots.).

LEMMA 2.3. *If* β_1 *and* $\beta_2 \in D$, *then we can find* $\alpha_0, \alpha_1, \cdots, \alpha_l$ *in* D *such that* $\pi(\alpha_0) = \pi(\beta_1)$ *and* $\pi(\alpha_l) = \pi(\beta_2)$ *and such that* $\sum_{i=0}^{j} \alpha_i$ *is a root for* $j = 0, 1, \cdots, l$.

Proof. The roots of D form a Dynkin diagram, and if this diagram is connected our assertion is clear. (We may take $\beta_1 = \alpha_0$ and $\beta_2 = \alpha_l$.) If D is not connected then the non-compact fundamental root occurs in the "interior" of the chain of the Dynkin diagram of $\mathfrak{g}$. Using Lemma 2.1 one can see that $\mathfrak{g} = A_n$ and we can check the lemma very simply in this case.

LEMMA 2.4. *For each* $i = 1, 2, \cdots, m$, *we can find* $\alpha \in D$ *such that* $(\pi(\alpha), \gamma_i) = (\alpha, \gamma_i) \neq 0$.

Proof. If $(\gamma_l, \alpha) = 0$ for all $\alpha \in D$, and if $l > 1$, $(\gamma_l, \gamma_1) = 0$. Now $D \cup \{\gamma_1\}$ is a fundamental system of roots and hence a basis in the Cartan subalgebra $\mathfrak{a}$, and γ_l is orthogonal to each one of these vectors. This is of course impossible. If $l = 1$, then we can find $\alpha \in D$ so that $(\gamma_1, \alpha) < 0$ since the Dynkin diagram of $\mathfrak{g}$ is connected.

We first examine those compact roots α for which $\pi(\alpha) = -\frac{1}{2}\gamma_i$.

LEMMA 2.5. *If* $\beta \in D$ *and* $\pi(\beta) = -\frac{1}{2}\gamma_i$, *then* $i = m$ (γ_m *is the highest strongly orthogonal root*).

Proof. First suppose that β_1 and β_2 are in D, and $\pi(\beta_1) = -\frac{1}{2}\gamma_i$, $\pi(\beta_2) = -\frac{1}{2}\gamma_k$ with $k > i$. We find a chain $\alpha_0, \alpha_1, \cdots, \alpha_l$ by Lemma 2.3.

We contend first that if λ_1 and λ_2 are compact positive roots so that $\lambda_1 + \lambda_2$ is a root, and if $\pi(\lambda_1) = -\frac{1}{2}\gamma_i$, then $\pi(\lambda_2) = \frac{1}{2}(\gamma_i - \gamma_j)$ or 0 for $j < i$ and $\pi(\lambda_1 + \lambda_2) = -\frac{1}{2}\gamma_j$, for $j \leqq i$. Suppose first that $\pi(\lambda_2) = -\frac{1}{2}\gamma_k$, then $\pi(\lambda_1 + \lambda_2) = -\frac{1}{2}\gamma_i - \frac{1}{2}\gamma_k$ a contradiction to Lemma 2.2. Then if $\pi(\lambda_2) \neq 0$, $\pi(\lambda_2) = \frac{1}{2}(\gamma_l - \gamma_j)$ for $j < l$. Since $\lambda_1 + \lambda_2$ is a root it follows from Lemma 2.2 that $l = i$, and we are done.

Since $\pi(\alpha_0) = -\frac{1}{2}\gamma_i$, we can argue inductively using the result of the previous paragraph to show that $\pi(\sum_{i=0}^{n} \alpha_i) = -\frac{1}{2}\gamma_j$ with $j \leqq i$ and that $\pi(\alpha_n) = 0$ or $\pi(\alpha_n) = -\frac{1}{2}(\gamma_s - \gamma_j)$, $j < s \leqq i$, for all $n = 1, 2, \cdots, l$. But for $n = l$ we find a contradiction since $\pi(\alpha_l) = -\frac{1}{2}\gamma_k$. Therefore among the $\pi(\alpha)$, $\alpha \in D$ at most one $-\frac{1}{2}\gamma_i$ occurs.

Suppose now $\beta_1 \in D$ and $\pi(\beta_1) = -\frac{1}{2}\gamma_i$ with $i < m$. If $\alpha \in D$, then $\pi(\alpha) \neq -\frac{1}{2}\gamma_m$ by the first part of the proof, but there exists $\beta_2 \in D$ so that $(\beta_2, \gamma_m) \neq 0$ by Lemma 2.4. It follows that $\pi(\beta_2) = \frac{1}{2}(\gamma_m - \gamma_n)$ for some

$n < m$. Now we choose $\alpha_0, \alpha_1, \cdots, \alpha_l$ according to Lemma 2.3. By the same argument we used in the first part of the proof for $k = 1, 2, \cdots, l$, $\pi(\alpha_k) = 0$ or $\pi(\alpha_k) = \frac{1}{2}(\gamma_s - \gamma_j)$ with $j < s \leqq i$, and in particular $s < m$. If we take $k = l$, $\pi(\alpha_l) = \frac{1}{2}(\gamma_m - \gamma_n)$ and we have a contradiction. Thus our Lemma is established.

Lemma 2.6. *For any $l < m$, there is an $\alpha \in D$ such that $\pi(\alpha) = \frac{1}{2}(\gamma_k - \gamma_l)$ $(k > l)$.*

Proof. If this were false for some $l < m$, then for all $\alpha \in D$, $\pi(\alpha) = \frac{1}{2}(\gamma_j - \gamma_i)$ with $j \neq l$ or $\pi(\alpha) = -\frac{1}{2}\gamma_m$. But by Lemma 2.3 $\gamma_l + \alpha$ is never a root. Moreover $\gamma_l + \gamma_1$ is not a root, and since $D \cup \{\gamma_1\}$ is a fundamental system of roots $\gamma_l + \beta$ is not a root for any fundamental root β. In this case γ_l must be the highest root of g, and since $\gamma_l < \gamma_m$, this is impossible. The lemma is therefore proved.

Lemma 2.7. *If β_1 and β_2 are in D, it is impossible that $\pi(\beta_1) = \frac{1}{2}(\gamma_k - \gamma_i)$ and $\pi(\beta_2) = \frac{1}{2}(\gamma_k - \gamma_j)$ with $i < j < k$.*

Proof. If this were the case, we may choose $\alpha_0, \cdots, \alpha_l$ by Lemma 2.3. We first suppose that for no r do we have $\pi(\alpha_r) = -\frac{1}{2}\gamma_m$. Then we claim that for $n = 0, 1, \cdots, l$, $\pi(\sum_{r=0}^{n} \alpha_r) = \frac{1}{2}(\gamma_t - \gamma_s)$ with $t \geqq k$ and $s \leqq i$. This is true for $n = 0$ and we proceed by induction. If it is true for n and $\pi(\alpha_{n+1}) = 0$, it is true for $n + 1$. Thus we may suppose that $\pi(\alpha_{n+1}) = \frac{1}{2}(\gamma_p - \gamma_q)$, and we have $\lambda = (\sum_{r=0}^{n+1} \alpha_r) = \frac{1}{2}(\gamma_t - \gamma_s) + \frac{1}{2}(\gamma_p - \gamma_q)$. In order not to contradict Lemma 2.3, we must have either $p = s$ or $t = q$. If $p = s$, then $q < p = s$ and $\lambda = \frac{1}{2}(\gamma_t - \gamma_q)$ with $t \geqq k$ and $q < s \leqq i$ as desired. If $t = q$, then $p > q = t$ and $\lambda = \frac{1}{2}(\gamma_p - \gamma_s)$ with $p > t \geqq k$ as desired. This also proves that if $\pi(\alpha_{n+1}) \neq 0$, then $\pi(\alpha_{n+1}) = \frac{1}{2}(\gamma_p - \gamma_q)$ with either $p > k$ or $q < i$, for $n = 0, 1, \cdots, l - 1$. Now we take $n = l - 1$, and $\pi(\alpha_l) = \frac{1}{2}(\gamma_k - \gamma_j)$ where we must have $j < i$, a clear contradiction.

If now $\frac{1}{2}(\alpha_{a+1}) = -\frac{1}{2}\gamma_m$ for some a, we may take a minimal. Then by the previous argument $\pi(\sum_{r=0}^{a} \alpha_r) = \frac{1}{2}(\gamma_t - \gamma_s)$ with $s \leqq i$. It is clear that we must have $t = m$ and that $\pi(\sum_{r=0}^{a+1} \alpha_r) = -\frac{1}{2}\gamma_s$ with $s \leqq i$. It follows by the argument given in the first part of Lemma 2.5 that for $r > a + 1$, $\pi(\alpha_r) = 0$ or $\pi(\alpha_r) = \frac{1}{2}(\gamma_p - \gamma_q)$ with $q < p \leqq i$. If we take $r = l$ then $\pi(\alpha_l) = \frac{1}{2}(\gamma_k - \gamma_j)$ with $i < j < k$, and this contradiction establishes the lemma.

We obtain our first main result.

THEOREM 1. *The non zero vectors in* $\pi(D)$ *are*

$$\tfrac{1}{2}(\gamma_2-\gamma_1),\cdots,\tfrac{1}{2}(\gamma_m-\gamma_{m-1})$$

together with perhaps $-\frac{1}{2}\gamma_m$.

Proof. In Lemma 2.6 we choose $l=m-1$, and it follows that $\frac{1}{2}(\gamma_m-\gamma_{m-1})$ is in $\pi(D)$. Now suppose that $\frac{1}{2}(\gamma_{i+1}-\gamma_i)$ is not in $\pi(D)$. We now choose a maximal such i, $(i<m-1)$. Now in Lemma 2.6 let $l=i$ and we find that $\pi(D)$ contains $\frac{1}{2}(\gamma_k-\gamma_i)$ for some $k>i+1$. By the maximality of i, $\pi(D)$ contains $\frac{1}{2}(\gamma_k-\gamma_{k-1})$ where $k-1>i$. Now if we let $j=k-1$, Lemma 2.7 is violated. Therefore $\pi(D)$ contains all vectors of the form $\frac{1}{2}(\gamma_{j+1}-\gamma_j)$ $j=1,2,\cdots,m-1$. Moreover if $\pi(D)$ contained some $\frac{1}{2}(\gamma_k-\gamma_i)$ with $k>i+1$, Lemma 2.7 is violated again. By Lemmas 2.2 and 2.5, $-\frac{1}{2}\gamma_m$ is the only other non zero vector which could belong to $\pi(D)$.

We denote by R the set of all roots of $\mathfrak{g}$.

THEOREM 2. 1) *There are two possibilities for* $\pi(R)$ *except for* 0.

(a) $\pi(R)\cup\{0\}=\{\pm\frac{1}{2}\gamma_i\pm\frac{1}{2}\gamma_j$ *all* i *and* $j\}$

(b) $\pi(R)\cup\{0\}=\{\pm\frac{1}{2}\gamma_i\pm\frac{1}{2}\gamma_j,\pm\frac{1}{2}\gamma_i$ *all* i *and* $j\}$

2) *The* γ_i *all have the same length.*

Proof. 1) We first claim that every $\frac{1}{2}(\gamma_j-\gamma_i)$ $j>i$ is contained in $\pi(C^+)$ (C^+ denotes the positive compact roots.) In fact we can find β_1 and β_2 in D such that $\pi(\beta_1)=\frac{1}{2}(\gamma_{i+1}-\gamma_i)$ and $\pi(\beta_2)=\frac{1}{2}(\gamma_j-\gamma_{j-1})$. (We are assuming as we may that $j>i+1$.) Then if we take $\alpha_0,\cdots,\alpha_l$ as in Lemma 2.3, it is easy to see that $\pi(\sum_0^l\alpha_r)=\frac{1}{2}(\gamma_j-\gamma_i)$.

Now suppose that $-\frac{1}{2}\gamma_m\notin\pi(D)$; then we claim that $-\frac{1}{2}\gamma_j\notin\pi(C^+)$. If $D=\{\mu_2,\cdots,\mu_n\}$, and $\alpha\in C^+$, $\alpha=\sum n_i\mu_i$. We shall show that $\pi(\alpha)\neq-\frac{1}{2}\gamma_j$ by induction on $l(\alpha)=\sum n_i$. This is true by Theorem 1 if $l(\alpha)=1$. If $l(\alpha)=l+1$ there is some i such that $\alpha-\mu_i$ is a root. Now $\pi(\alpha-\mu_i)=\frac{1}{2}(\gamma_k-\gamma_l)$, and since $\pi(\mu_i)=\frac{1}{2}(\gamma_{r+1}-\gamma_r)$ or 0, it is clear that $\pi(\alpha)\neq-\frac{1}{2}\gamma_j$ for any j.

Combining these results with those in [5], p. 587, we see that if $-\frac{1}{2}\gamma_m\notin\pi(D)$ then we have the first alternative for $\pi(R)$.

Now suppose that $-\frac{1}{2}\gamma_m\in\pi(D)$. In order to prove that $\pi(R)$ is given by the second alternative, all we need show is that $-\frac{1}{2}\gamma_i \in \pi(C^+)$ for any i. But if we take according to Lemma 2.3 $\alpha_0, \cdots, \alpha_l$ such that $\pi(\alpha_0) = -\frac{1}{2}\gamma_m$ and $\pi(\alpha_l) = \frac{1}{2}(\gamma_{i+1} - \gamma_i)$ it is clear that $\pi(\sum_{a=0}^{l} \alpha_l) = -\frac{1}{2}\gamma_i$. The first part of the theorem is now proved.

2) If $i < j$, then we can find $\alpha \in C^+$ such that $\pi(\alpha) = \frac{1}{2}(\gamma_j - \gamma_i)$. Since $(\alpha, \gamma_i) = -\frac{1}{2}(\gamma_i, \gamma_i) < 0$, $\gamma_i + \alpha$ is a root. Suppose that $\alpha \neq \pi(\alpha)$. Then we claim that $\lambda = \gamma_i + 2\alpha$ is not a root. If it were, it would be positive and non-compact with $\pi(\beta) = \gamma_j$. Then $(\beta, \gamma_j) = (\gamma_j, \gamma_j) > 0$, so that $\lambda = \beta - \gamma_j$ is a root. But now λ is compact and non-zero since $\alpha \neq \pi(\alpha)$. On the other hand $\pi(\lambda) = 0$, and by Lemma 2.2, λ is strongly orthogonal to γ_j. Since $\beta = \lambda + \gamma_i$ is a root, we have a contradiction so that $\gamma_i + 2\alpha$ is not a root. But now $\gamma_i - \alpha$ is not a root since $(\gamma_i, \alpha) < 0$ and the α chain through γ_i has length at most 4 ([9], p. 117). It follows that $2(\gamma_i, \alpha)/(\alpha, \alpha) = -1$.

We know that $\gamma_j + \alpha$ is not a root (Lemma 2.2), and since $(\gamma_j, \alpha) > 0$, $\gamma_j - \alpha$ is a root. By the same argument as above, it follows that $\gamma_j - 2\alpha$ is not a root, and we can conclude that $2(\gamma_j, \alpha)/(\alpha, \alpha) = 1$. Combining these results we see that $(\gamma_j, \alpha) = -(\gamma_i, \alpha)$ or that $\frac{1}{2}(\gamma_j, \gamma_j) = \frac{1}{2}(\gamma_i, \gamma_i)$ as desired.

Now let us suppose that $\alpha = \pi(\alpha) = \frac{1}{2}(\gamma_j - \gamma_i)$. Then $\gamma_i + 2\alpha = \gamma_j$ is a root, whereas $\pi(\gamma_i - \alpha) = \frac{3}{2}\gamma_i - \frac{1}{2}\gamma_j$ and $\pi(\gamma_i + 3\alpha) = \frac{3}{2}\gamma_j - \frac{1}{2}\gamma_i$ so that $\gamma_i - \alpha$ and $\gamma_i + 3\alpha$ are not roots. Thus $2(\alpha, \gamma_i)/(\alpha, \alpha) = -2$. By the same argument, $2(\alpha, \gamma_j)/(\alpha, \alpha) = 2$ and we find as before that $(\gamma_j, \gamma_j) = (\gamma_i, \gamma_i)$.

Remark. Koranyi and Wolf, [11] have shown that alternative 1 is a necessary and sufficient condition that G/K be realizable as a tube domain over a domain of positivity.

Section 3. We shall now apply these results to study the Cartan domains G/K. We observe (cf. [5], p. 219) that the root vectors X_α, $\alpha \in R$ may be normalized so that $\tau(X_\alpha) = -X_{-\alpha}$ where τ is the involution of $\mathfrak{g}$ leaving $\mathfrak{u}$ fixed. Also we choose these so that $[[X_\alpha, X_{-\alpha}], X_\alpha] = 2X_\alpha$ and so that $N_{\alpha,\beta}$ is real and satisfies $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$ where $[X_\alpha, X_\beta] = N_{\alpha,\beta}X_{\alpha+\beta}$ if $\alpha + \beta$ is a root. We denote by λ_i the vectors $(X_{\gamma_i} + X_{-\gamma_i})/2$ where the X_α are fixed once and for all. Harish-Chandra's fundamental result is the following ([5], p. 314)

THEOREM. *The vectors $\lambda_1, \cdots, \lambda_m$ span a maximal abelian subspace of $\mathfrak{p}_0$.*

Now we write $\mathfrak{a}_0 = \mathfrak{a}_0^+ + \mathfrak{a}_0^-$ where $i\mathfrak{a}_0^-$ is the (real) subspace spanned by γ_i, and $\mathfrak{a}_0^+$ is the orthogonal complement. Further let c be the Cayley transform of $\mathfrak{g}$ defined by

$$c = \mathrm{ad}(\exp(\pi/4)\sum_{i=1}^{m}(X_{-\gamma_i} - X_{\gamma_i})).$$

This is essentially the Cayley transform in [11]. Then c is the identity transformation on $\mathfrak{a}_0^+$ and $c([X_{\gamma_i}, X_{-\gamma_i}]) = 2\lambda_i$. Since $[X_{\gamma_i}, X_{-\gamma_i}] = 2/d\gamma_i$ where d is the common value of (γ_i, γ_i), $c(i\mathfrak{a}_0^-)$ is a maximal abelian subspace of $\mathfrak{p}_0$. Then $c(\mathfrak{a}_0^+ + i\mathfrak{a}_0^-) = \mathfrak{a}_0^+ + \mathfrak{h}^- = \mathfrak{h}_0$ is a Cartan subalgebra of $\mathfrak{g}_0$ such that $\mathfrak{h}_0 = (\mathfrak{h}_0 \cap \mathfrak{k}_0) + (\mathfrak{h}_0 \cap \mathfrak{p}_0)$ and $\mathfrak{h}_0 \cap \mathfrak{p}_0$ is maximal abelian in $\mathfrak{p}_0$. If $\mathfrak{h}$ is the complexification of $\mathfrak{h}_0$, $c(\mathfrak{a}) = \mathfrak{h}$ so that set of roots Σ of $\mathfrak{g}$ with respect to $\mathfrak{h}$ is $c(R)$ where R is the set of roots of $\mathfrak{g}$ with respect to $\mathfrak{a}$. We can rephrase the result of Section 2 in this terminology. (We use π again to denote the projection from $\mathfrak{h}_0$ to $\mathfrak{h}^-$.

THEOREM 2′. *The set of projections of Σ into $\mathfrak{h}^-$ has one of the two forms*

1.) $\pi(\Sigma) \cup \{0\} = \{d(\pm\tfrac{1}{2}\lambda_i \pm \tfrac{1}{2}\lambda_j)$ *all i and j*$\}$

2.) $\pi(\Sigma) \cup \{0\} = \{d(\pm\tfrac{1}{2}\lambda_i \pm \tfrac{1}{2}\lambda_j), d(\pm\tfrac{1}{2}\lambda_i) \cdots$ *all i and j*$\}$.

We totally order the real vector space $\mathfrak{h}^-$ by the relation $x > 0$ if $x = \sum a_i\lambda_i$ and the last non-zero a_i is positive. We totally order $i\mathfrak{a}_0^+$ and define a total ordering on $\mathfrak{h}^- + i\mathfrak{a}_0^+$ by $(h^-, h^+) > 0$ if $h^- > 0$ or if $h^- = 0$, then $h^+ > 0$. Let $\alpha_1, \cdots, \alpha_n$ be a fundamental system of roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$ in this ordering. Then we let F be the set of distinct non-zero restrictions to $\mathfrak{h}^-$ of the (α_i) (F is called a restricted fundamental system of $\mathfrak{g}_0$, [15]). Finally let W be the restricted Weyl group of $\mathfrak{g}_0$ and Λ be the open positive Weyl Chamber in $\mathfrak{h}^-$.

THEOREM 3. 1) *F has one of the two forms*

a) $d\lambda_1, (d/2)(\lambda_2 - \lambda_1), \cdots, (d/2)(\lambda_m - \lambda_{m-1})$

b) $(d/2)\lambda_1, (d/2)(\lambda_2 - \lambda_1), \cdots, (d/2)(\lambda_m - \lambda_{m-1})$.

2) *W consists of all transformations of the form $\lambda_i \to \pm\lambda_{\sigma(i)}$ where σ is a arbitrary permutation of $1, \cdots, m$.*

3) $\Lambda = \{H = \sum a_i\lambda_i : 0 < a_1 < a_2 < \cdots < a_n\}$.

Proof. We note that the set of non-zero restrictions to $\mathfrak{h}^-$ of the positive roots of g with respect to $\mathfrak{h}$ are $(d/2)(\gamma_j - \gamma_i)$, $j > i$, $(d/2)(\gamma_j + \gamma_i)$ and $d/2\,\gamma_j$ if the latter occur at all. The set F may be characterized as the smallest subset of these such that each one can be written as an non-negative integral linear combination of vectors in F ([15], p. 81). If no $d/2\,\gamma_i$ occurs, this set is clearly given by (a) and if these do occur it is clearly given by (b). This proves the first statement.

It is known ([15], p. 81) that the set F forms a Dynkin diagram and that W and Λ are respectively the Weyl group and open positive chamber of this diagram. We observe that in (a) F is the diagram of C_m and in (b) F is the diagram of B_m. In both cases the Weyl group and positive chamber are given by (2) and (3) respectively.

According to [4] G/K may be embedded holomorphically as a bounded domain in the complex vector space $\mathfrak{p}^-$. The adjoint representation of $\mathfrak{k}_0$ on $\mathfrak{p}^-$ is equivalent to the adjoint representation of $\mathfrak{k}_0$ in $\mathfrak{p}_0$. Thus since $\mathfrak{p}^-$ is irreducible, $\mathfrak{p}_0$ can be uniquely given the structure of a complex vector space so that these representations are equivalent as complex representations. On can check, using the normalization of the X_α that $X_{-\beta} \rightarrow (X_\beta + X_{-\beta})/2$ and $iX_{-\beta} \rightarrow i(X_{-\beta} - X_\beta)/2$ for $\beta \in P^+$ is a k_0-invariant map over R. It will be simpler to use $\mathfrak{p}_0$ with this complex structure in place of $\mathfrak{p}^-$. Using [4], we see that G/K may be embedded as a bounded domain and the mapping which we denote by z is given by

$$z(k \cdot \exp(\sum a_i\lambda_i)K) = \operatorname{ad}(k)(\sum \tanh(a_i/2)\lambda_i)$$

where $\sum a_i\lambda_i \in \mathfrak{h}^-$ and $k \in K$. Let $\mathfrak{D} = z(G/K)$ and $\bar{\mathfrak{D}}$ the closure of $\mathfrak{D}$. Then it can be seen that the action of G extends continuously to $\bar{\mathfrak{D}}$ so that $\bar{\mathfrak{D}}$ is a compactification of $\mathfrak{D}$. We shall now show that it is one of the Satake-Furstenberg compactifications.

These compactifications are in one-to-one correspondence with faithful ([13], § 4) subsets of F (F is the restricted fundamental system of $\mathfrak{g}_0$.) Since $\mathfrak{g}_0$ is simple, all subsets except F itself are faithful. If $E \subset F$, then we recall from [13] that $\overline{\phi_E(G/K)}$ is the corresponding compactification of G/K.

THEOREM 4. *If* $E = \{d/2(\lambda_2 - \lambda_1), \cdots, d/2(\lambda_m - \lambda_{m-1})\}$ *then* $\overline{\phi_E(G/K)}$ *is homeomorphic as G-space to* $\bar{\mathfrak{D}}$. *The subset of point masses,* $G/B(E)$ *in* $\overline{\phi_E(G/K)}$ *is homeomorphic to the Bergmann-Shilov boundary of* $\mathfrak{D}$.

The first step in the proof will be to identify the Bergmann-Shilov boundary of D. Let $M_0 = \overline{\phi(G/K)}$ be the maximal compactification defined by Furstenberg; the corresponding subset of F is the void set. Then Furstenberg [2], p. 370 shows that one can map M_0 equivariantly onto $\bar{\mathfrak{D}}$. We briefly recall this construction. If z_i is any coordinate function on $\mathfrak{p}_0$, z_i is analytic and since $\mathfrak{D}$ is Kähler, z_i is a harmonic function which is continuous on $\bar{\mathfrak{D}}$. By Furstenberg's boundary value theorem z_i has a boundary value f_i

which in this case is a continuous function on G/B. Now $G/B \subset M_0$ as the set of point masses and f_i can be extended continuously to all of M_0 by

$$f_i(\mu) = \int f_i(b)\, d\mu(b)$$

for $\mu \in M_0$ (of course if $\mu = \phi(s)$, $s \in G/K$, then $f_i(\mu) = z_i(z(s))$ where z is the embedding of G/K into $\mathfrak{p}_0$.) Now he defines a map f of M_0 into $\mathfrak{p}_0$ by

$$f(\mu) = (f_1(\mu), \cdots, f_n(\mu))$$

where $z_1, \cdots, z_n$ are complex coordinates on $\mathfrak{p}_0$. This is clearly independent of the coordinates, continuous and equivariant, and maps M_0 onto $\bar{\mathfrak{D}}$. He also shows that $f(G/B)$ is the Shilov boundary S of $\mathfrak{D}$. Thus it is clear that f sets up a homeomorphism of G/C with S where C is some subgroup containing B.

LEMMA 3.1. *If E is as in the statement of Theorem 4, then $S \cong G/B(E)$.*

Proof. We first claim that $f(B) = \sum_{i=1}^{m} \lambda_i = \epsilon \in \bar{D}$, where f is the equivariant map of G/B onto S. Let H be a regular element of $\Lambda \subset \mathfrak{h}^-$; that is $H = \sum a_i\lambda_i$ with $0 < a_1 < a_2 < \cdots < a_m$. If $b_t = \exp(tH)$, we know from [13] that $\phi(b_tK)$ tends to a point mass at B in G/B. Therefore $f(B) = \lim_{t\to\infty} z(b_tK)$ where the limit is taken in $\bar{\mathfrak{D}}$. But clearly making use of the formula for z above, we see that $f(B) = \epsilon$ as desired.

Now as remarked above $S \cong G/C$ where $C \supset B$ and by the previous paragraph C is the subgroup of G which leaves $\epsilon \in \bar{\mathfrak{D}}$ fixed. We know that $C = B(D)$ for some $D \subset F$, and by the above $C \cap K$ must be the centralizer of ϵ in K. Also $C \cap K = B(D) \cap K = M(D)$. Therefore $D = \{\lambda \colon \lambda \in F \colon \lambda(\epsilon) = 0\}$ and it is clear that this is precisely the set E described in Theorem 4.

We can now construct a map from $M = \overline{\phi_B(G/K)}$ onto $\bar{\mathfrak{D}}$. If h is the projection mapping of G/B onto $G/B(E)$, and μ is a probability measure on G/B, $h(\mu)$ is the probability measure on $G/B(E)$ defined by $h(\mu)(A) = \mu(h^{-1}(A))$ for a Borel subset A of $G/B(E)$. It is easy to see that h maps $M_0 = \overline{\phi(G/K)}$ equivariantly onto $M = \overline{\phi_E(G/K)}$. Recall that the map f of M onto $\bar{\mathfrak{D}}$ defined above has the property proved in Lemma 3.1 that $f(\delta_1) = f(\delta_2)$ if δ_1 and δ_2 are point masses at b_1 and b_2 in G/B such that $h(b_1) = h(b_2)$. It follows from the definition of f that $f(\mu_1) = f(\mu_2)$ if $\mu_i \in M_0$ and $h(\mu_1) = h(\mu_2)$. Therefore we can define a map g from M onto $\bar{\mathfrak{D}}$

by $g(\mu) = f(\mu')$ when $\mu' \in M_0$ and $h(\mu) = \mu'$. It is clear that g is continuous and equivariant. It is easy to prove the following characterization of g.

LEMMA 3.2. *If b is any harmonic function on $\mathfrak{D} \cong G/K$, continuous on $\bar{\mathfrak{D}}$, then b restricted to S corresponds to the abstract boundary function on $G/B(E)$ and*

$$b(g(\mu)) = \int_{G/B(E)} b(s)\, d\mu(s)$$

for any $\mu \in M$.

The remainder of this section is devoted to the proof that g is one-to-one. For this we will need to use the detailed structure of M given in [13].

LEMMA 3.3. *If $E = \{d/2(\lambda_2 - \lambda_1), \cdots, d/2(\lambda_m - \lambda_{m-1})\}$ then the E-connected subsets (cf. [13]) are of the form*

$$D_k = \{d\lambda_1(d/2\,\lambda_1), d/2(\lambda_i - \lambda_{i-1}) \;\; 1 = 2, \cdots, k\}$$

for $k = 0, 1, 2, \cdots, m$, where we understand D_0 to be void.

Proof. Clearly each D_k is E-connected since if $k \neq 0$, all elements of D_k are connected to $d\lambda_1(d/2\lambda_1)$. Let D be a non void E-connected set; then if $D \subset E$ it is clearly not E-connected. Thus $d\lambda_1(d/2\lambda_1)$ lies in D. Let D_k be the largest set of this form contained in D. If $D \neq D_k$, then $d/2(\lambda_{k+1} - \lambda_k) \notin D$ and $D - D_k$ is clearly a union of components of D which lie entirely in E. Thus, D is not E-connected by definition, and we are done.

LEMMA 3.4. 1) *If ν_i $i = 1, 2$ belongs to $\overline{\phi_B(\exp(\bar{\Lambda})K)}$ and if $g(\nu_1) = g(\nu_2)$, then $\nu_1 = \nu_2$.*

2) *If $\nu \in \overline{\phi_B(\exp(\bar{\Lambda})K)}$, then we can find a sequence $H_n \in \bar{\Lambda}$ such that $\phi_B(\exp(H_n)K) \to \nu$ and such that $\mathrm{ad}(k)H_n = H_n$ for any k such that $\mathrm{ad}(k)g(\nu) = g(\nu)$.*

Proof. (1) Let μ_k be the $K(D_k)$ invariant measure on the orbit $O(E)^{D_k}$. Now let $H^n \in \bar{\Lambda}$ so that $\nu = \lim \phi_B(\exp(H^n)K)$ exists. Then $H^n = \sum a_i^n \lambda_i$ with $0 \leqq a_1^n \leqq a_2^n \leqq \cdots \leqq a_m^n$ for each n. We may suppose that $\lambda(H^n)$ has a non negative limit (finite or $+\infty$) for each $\lambda \in F$. Let D be the subset of all $\lambda \in F$ such that $\lambda(H^n)$ has a finite limit, and let D_k be the maximal E-connected subset of D. Then it is clear that $a_i^n \to a_i^0$ for $i = 1, 2, \cdots$, and that $a_{k+i}^n \to +\infty$ for $i \geqq 1$. If we define H^0 by $\sum_{i=1}^{k} a_i^0 \lambda_i$, it follows from [13], Theorem 7 that $\nu = \exp(H^0) \cdot \mu_k$. Now

$$g(\nu) = \lim \sum_{i=1}^{m} \tanh(a_i^n/2)\lambda_i)$$

$$= \sum_{i=1}^{k} \tanh(a_i^0/2)\lambda_i + \sum_{k+1}^{m} \lambda_i.$$

Thus if we are given $g(\nu) = \sum t_i\lambda_i$, we can determine the index k by the rule that t_{k+1} is the first coefficient equal to 1. Further the t_i, $i \leqq k$ determine H^0 completely so that $g(\nu)$ will determine $\nu = \exp(H^0)\cdot\mu_k$. This proves part (1) of the Lemma.

(2) Let $Y = g(\nu)$ be given. Then if $D = \{\lambda\colon \lambda \in F, \lambda(Y) = 0\}$, we see that the centralizer of Y in K is $M(D)$. Conversely the centralizer of $M(D)$ in $\mathfrak{h}^-$ is $\{H\colon \lambda(H) = 0 \text{ for all } \lambda \in D\}$. It is now completely obvious that one can choose a suitable sequence H^n.

We shall now show that the map g is one-to-one. Suppose that $\mu_i \in M$ $i = 1, 2$, and that $g(\mu_1) = g(\mu_2)$. We know from [13] that we can find ν_1 and ν_2 in $\phi_B(\exp(\bar{\Lambda})K)$ and k_1 and k_2 in K so that $\mu_i = k_i\nu_i$. If $Y_i = g(\nu_i)$ then $g(\mu_1) = g(\mu_2)$ means that $\mathrm{ad}(k_1)Y_1 = \mathrm{ad}(k_2)Y_2$ or that $Y_2 = \mathrm{ad}(k)Y_1$ where $k = k_2^{-1}k_1$. Since $Y_i \in \mathfrak{h}^-$ we know that $k = s\cdot m$ where m centralizes Y_i and s is in the restricted Weyl group. Thus $Y_2 = \mathrm{ad}(s)Y_1$ and since Y_1 and Y_2 both lie in the closed positive chamber $\bar{\Lambda}$ it is easy to see that $Y_1 = Y_2$ and that k centralizes Y_1. By part (1) of Lemma 3.4, $\nu_1 = \nu_2$.

Furthermore $\nu_1 = \lim \exp(H^n)\mu$ where μ is the K invariant measure on $S \cong G/B(E)$. In virtue of part (2) of Lemma 3.4, we may also assume that $\mathrm{ad}(k)H^n = H^n$. Thus

$$k\cdot\nu_1 = \lim \exp(\mathrm{ad}(k)H^n)k\mu = \nu_1$$

since $k\cdot\mu = \mu$. Finally

$$\mu_2 = k_2\nu_2 = k_2\nu_1 = k_1k_1^{-1}k_2\nu_1 = k_1k^{-1}\nu_1 = k_1\nu_1 = \mu_1$$

so that the mapping g is on-to-one. Theorem 4 is completely proved.

We note that Theorem 4 gives us an affirmative solution to the Dirichlet problem for $\bar{\mathfrak{D}}$ relation to S cf. [6], [12]. A function f is harmonic in $\mathfrak{D}$ if $\Delta f = 0$ where Δ is the G-invariant Laplace operator on $\mathfrak{D}$.

THEOREM 5. *If b is any continuous function on the Shilov boundary S of $\mathfrak{D}$, then there is one and only one function continuous on $\bar{\mathfrak{D}}$, harmonic on $\mathfrak{D}$ and taking the given boundary values on S.*

Proof. If we replace $\bar{\mathfrak{D}}$ by $M = \overline{\phi_B(G/K)}$ and S by $G/B(E) \subset M$, the existence and uniqueness of the harmonic functions is clear from Furstenberg's work. Then Theorem 4 immediately implies the present theorem.

Section 4. In this section we shall treat Pyatetskiĭ-Shapiro's [14] theory of boundary components using the tools we have developed in the first three sections. We preserve our notation from the previous section. Let $D = D_k$ be an E-connected set, and we consider the corresponding subalgebra $\mathfrak{g}(D)$ of $\mathfrak{g}$ defined in [13].

As before let $c = \operatorname{ad}(\exp(\pi/4 \sum_{i=1}^{m} (X_{-\gamma_i} - X_{\gamma_i}))$ be the Cayley transform of $\mathfrak{g}$ and let $c_D = \operatorname{ad}(\exp(\pi/4 \sum_{i=1}^{k} (X_{-\gamma_i} - X_{\gamma_i}))$. We know that $c(\mathfrak{a}) = \mathfrak{h}$ and also $c(\gamma_i) = d\lambda_i$ $i = 1, \cdots, m$. If $\mathfrak{g}'(D)$ is the derived algebra of the centralizer of the set of elements $\gamma_{k+1}, \cdots, \gamma_m$, then it is clear from the definition of $\mathfrak{g}(D)$ that $c(\mathfrak{g}'(D)) = \mathfrak{g}(D)$. On the other hand

$$\mathfrak{g}'(D) = (\mathfrak{g}'(D) \cap \mathfrak{a}^+) + \sum_{i=1}^{k} C \cdot \gamma_i + \sum C \cdot X_\beta \quad \beta \in S$$

where S is the set of roots of $\mathfrak{g}$ with respect to $\mathfrak{a}$ such that $(\beta, \gamma_i) = 0$ $i = k+1, \cdots, m$. But if $(\beta, \gamma_i) = 0$ we know that β is strongly orthogonal to γ_i and hence $c(X_\beta) = c_D(X_\beta)$ $\beta \in S$. Further $c = c_D = 1$ on $\mathfrak{a}^+$ and $c(\gamma_i) = c_D(\gamma_i)$ for $i = 1, \cdots, k$. Thus $c = c_D$ on $\mathfrak{g}'(D)$. Finally we observe that $\pm\gamma_i \in S$ and hence $c_D \in \exp(\mathfrak{g}'(D))$. Therefore $\mathfrak{g}'(D) = c_D(\mathfrak{g}'(D)) = c(\mathfrak{g}'(D)) = \mathfrak{g}(D)$. We have proved the following facts.

LEMMA 4.1. $\mathfrak{g}(D) = \mathfrak{g}'(D)$ *and* $\mathfrak{g}(D)$ *is invariant under* c.

Recall that $G(D)$, the subgroup of G corresponding to the subalgebra $\mathfrak{g}(D) \cap \mathfrak{g}_0$ is semi-simple, and $G(D) \cap K = K(D)$ is a maximal compact subgroup.

LEMMA 4.2. *$G(D)/K(D)$ is an irreducible Hermitian symmetric space.*

Proof. We first observe that if $\mathfrak{p}_0(D)$ is the orthogonal complement of $\mathfrak{k}_0(D)$ in $\mathfrak{g}_0(D)$, then $\mathfrak{p}_0(D) \subset \mathfrak{p}_0$. Since $\mathfrak{h}^-(D) \subset \mathfrak{p}_0(D)$ and

$$\mathfrak{p}^0(D) = \operatorname{ad}(K(D)) \cdot \mathfrak{h}^-(D),$$

it is enough to observe that $\mathfrak{h}^-(D) \subset \mathfrak{p}_0$. Therefore the complexification of $\mathfrak{p}_0(D)$, $\mathfrak{p}(D)$ lies in $\mathfrak{p}$.

We know that D is a restricted fundamental system of roots of $\mathfrak{g}_0(D)$, ([15], p. 83). Since the Dynkin diagram of D is connected, it follows that $G(D)/K(D)$ is an irreducible symmetric space. Further it is known that it will be Hermitian symmetric if and only if the adjoint representation of $K(D)$ on $\mathfrak{p}(D)$ is not irreducible.

But if $\mathfrak{p} = \mathfrak{p}^{+} + \mathfrak{p}^{-}$ is the splitting of $\mathfrak{p}$ under $\mathrm{ad}(K)$, we know at once that $\mathfrak{p}^{\pm} \cap \mathfrak{p}(D)$ are invariant under $K(D)$. But now by Lemma 4.1, $X_{\pm\gamma_1} \in \mathfrak{p}^{\pm} \cap \mathfrak{p}(D)$ so that both of these spaces are non-trivial. Thus $\mathrm{ad}\, K(D)$ is not irreducible on $\mathfrak{p}(D)$ and the lemma is proved.

It is clear that $\mathfrak{a}(D) = \mathfrak{g}(D) = \mathfrak{g}(D) \cap \mathfrak{a}$ is a Cartan subalgebra of $\mathfrak{k}(D)$ and $\mathfrak{g}(D)$. Moreover the roots of $\mathfrak{g}(D)$ with respect to $\mathfrak{a}(D)$ are exactly the roots of $\mathfrak{g}$ which lie in $\mathfrak{a}(D)$. The roots of $\mathfrak{g}(D)$ can then be ordered compatibly with the roots of g. It is clear that any fundamental root of $\mathfrak{g}$ which is a root of $\mathfrak{g}(D)$ is a fundamental root of $\mathfrak{g}(D)$.

LEMMA 4.3. *The set* $\{\gamma_1, \gamma_2, \cdots, \gamma_k\}$ *are the Harish-Chandra strongly orthogonal roots of* $\mathfrak{g}(D)$.

Proof. By the comment above, γ_1 is a fundamental root of $\mathfrak{g}(D)$ and is non-compact since $X_{\gamma_1} \in \mathfrak{p}(D)$. Therefore it is the unique non-compact fundamental root of $\mathfrak{g}(D)$. Furthermore, γ_i is the smallest positive non-compact root of $\mathfrak{g}(D)$ strongly orthogonal to $\gamma_1, \cdots, \gamma_{i-1}$, $(i \leq k)$ and this suffices to characterize them as the strongly orthogonal roots of Harish-Chandra.

Remark. It follows at once that the mapping c_D defined in the proof of Lemma 4.1 is the Cayley transform of $g(D)$.

Let z_D be the Harish-Chandra embedding of $G(D)/K(D)$ as a bounded domain in $\mathfrak{p}_0(D)$. We note first that it is easy to see $\mathfrak{p}_0(D)$ with the complex structure it inherits, is a complex subspace of $\mathfrak{p}_0$. We can obtain a formula for z_D by making the appropriate changes in the formula for z in Section 3. We obtain the following:

$$z_D(k \exp(\sum_{i=1}^{k} a_i\lambda_i) K(D)) = \mathrm{ad}(k)(\sum_{i=1}^{k} \tanh(a_i/2)\lambda_i)$$

for $k \in K(D)$.

On the other hand it follows from [13] and Section 3 that the symmetric space $S_D = G(D)/K(D)$ may be embedded in the boundary $\partial\mathfrak{D}$ of $\mathfrak{D}$. Denote this mapping by i_D. Then if $k \in K(D)$,

$$i_D(k \exp(\sum_{i=1}^{k} a_i\lambda_i) K(D)) = \mathrm{ad}(k) \lim_{t\to\infty} z(\exp(tH_D + \sum_{i=1}^{k} a_i\lambda_i) K)$$

where $H_D = \lambda_{k+1} + 2\lambda_{k+2} + \cdots + (m-k)\lambda_m$ is a "regular" element of $\mathfrak{v}(D)$ the orthogonal complement of $\mathfrak{h}^{-}(D)$ in $\mathfrak{h}^{-}$. The limit can be evaluated at once, and it is equal to

$$\operatorname{ad}(k)\,(\sum_{i=1}^{k}\tanh(a_i/2)\lambda_i)+\operatorname{ad}(k)\epsilon_D$$

where $\epsilon_D=\sum_{i=k+1}^{m}\lambda_i$. Finally we note that ϵ_D is centralized by $K(D)$ so that $\operatorname{ad}(k)\epsilon_D=\epsilon_D$.

A comparison of the formula for z_D and i_D yields the following Lemma.

LEMMA 4.4. *For* $s\in S_D$, $i_D(s)=\epsilon_D+z_D(s)$.

The content of Theorem 7 of [13] is that the sets $k\cdot i_D(S_D)$ for all E-connected sets D unequal to F, and all $k\in K$, either coincide or are disjoint and exhaust the boundary $\partial(\mathfrak{D})$ of $\mathfrak{D}$. We shall show that these sets are exactly the boundary components of Pyatetskiĭ-Shapiro.

The following lemma will play a crucial role in the sequel. This result was obtained by R. Hermann previously (unpublished). We denote by τ the conjugate linear involution of $\mathfrak{g}$ whose fixed points are $\mathfrak{u}$. Then if $(\ ,\ )$ is the Killing form on $\mathfrak{g}$, $H(u,v)=-(u,\tau(v))$ defines a positive definite Hermitian form on $\mathfrak{g}$. Furthermore if $P\in\mathfrak{p}_0$, $\operatorname{ad}(P)$ is an Hermitian transformation with respect to this form. We use $|\operatorname{ad}(P)|$ to denote the norm of $\operatorname{ad}(P)$.

LEMMA 4.5. *The domain* $D=\{P: P\in\mathfrak{p}_0$ *and* $|\operatorname{ad}(P)|<1\}$.

Proof. If $k\in K$, $|\operatorname{ad}(\operatorname{ad}(k)\cdot P)|=|\operatorname{ad}(P)|$ since $\operatorname{ad}(k)$ is a unitary transformation. Since $\mathfrak{p}_0=\operatorname{ad}(K)\mathfrak{h}^-$ and $\mathfrak{D}=\operatorname{ad}(K)\ (\mathfrak{h}^-\cap\mathfrak{D})$ it is enough to show that

$$\mathfrak{D}\cap\mathfrak{h}^-=\{H: H\in\mathfrak{h}^-\,|\operatorname{ad}(H)|<1\}.$$

But if $H=\sum t_i\lambda_i$, then in a Weyl basis of $\mathfrak{g}$, $\operatorname{ad}(H)$ is diagonal with eigenvalues which can be read off from Theorem 2′. To be precise the eigenvalues are $dl/2(\pm t_i\pm t_j)$ together with perhaps $dl/2(\pm t_i)$ where l is the common value of (λ_i,λ_i), and d is the common value of (γ_i,γ_i). From the definitions we see that $c(\gamma_i)=d\lambda_i$ and since c preserves the Killing form $d=(\gamma_i,\gamma_i)=d^2(\lambda_i,\lambda_i)=d^2l$. Thus $dl=1$ and $\operatorname{ad}(H)$ has eigenvalues $1/2(\pm t_i\pm t_j)$ together with perhaps $\pm\frac{1}{2}t_i$. It follows that

$$\{H: |\operatorname{ad}(H)|<1\}=\{H: H=\sum t_i\lambda_i\,|\,t_i|<1\},$$

but the later set we know is just $\mathfrak{D}\cap\mathfrak{h}^-$.

Remark. It is interesting to note that the function $|\operatorname{ad}(P)|$ is plurisubharmonic (PSH). To see this, we note that $|\operatorname{ad}(P)|=\sup_{|x|=1}|H([P,x],x)|$, since $\operatorname{ad}(P)$ is Hermitian. Now $H([P,x],x)$ is a real linear functional on $\mathfrak{p}_0$, hence it is PSH. The absolute value of a PSH function is likewise PSH and the supremum of a family of PSH functions is PSH provided it is continuous as $|\operatorname{ad}(P)|$ is.

LEMMA 4.6. 1) *If* $x \in \bar{\mathfrak{D}}$ *then* $|(x,\lambda_i)| \leqq (\lambda_i,\lambda_i)$

2) *Let* $x = \epsilon_D + y \in \mathfrak{p}_0$ *where* $(y,\lambda_i) = 0$ *if* $i > k$. *Then if* $x \in \bar{\mathfrak{D}}$, *we must have* $y \in \mathfrak{p}_0(D)$.

Proof. We observe that the vectors

$$a_\beta = (X_\beta + X_{-\beta})/2 \text{ and } b_\beta = i/2(X_\beta - X_{-\beta})$$

for $\beta \in P^+$ form an orthogonal bases in $\mathfrak{p}_0$. Finally let $c_\beta = [X_\beta, X_{-\beta}]/2 = 1/(\beta,\beta)\beta$ (see [5], p. 315). Then we see that

$$[a_\beta, ib_\beta + c_\beta] = ib_\beta + c_\beta,$$
$$[a_\gamma, ib_\beta + c_\beta] = K + (\gamma,\beta)/(\beta,\beta) ib_\gamma \text{ if } \beta \neq \gamma,$$

where $K \in k$ and $K \perp \mathfrak{a}$,

$$[b_\gamma, ib_\beta + c_\beta] = (\gamma,\beta)/(\beta,\beta)(-ia_\gamma) + K', \; K' \text{ as above.}$$

For convenience, we let $z_i = ib_{\gamma_i} + c_{\gamma_i}$.

For part (1) of the lemma, let $x \in \bar{\mathfrak{D}}$ and $x = \sum t_\gamma a_\gamma + s_\gamma b_\gamma$ $(\gamma \in P^+)$. Then

$$[x, z_i] = t_{\gamma_i} z_i + \sum s_\gamma (\gamma,\gamma_i)/(\gamma_i,\gamma_i)(-ia_\gamma) + \sum_{\gamma \neq \gamma_i} t_\gamma (\gamma,\gamma_i)/(\gamma_i,\gamma_i) ib_\gamma + K$$

where $K \in k$ and $K \perp \mathfrak{a}$. In the Hermitian metric on $\mathfrak{g}$ all terms but the first are perpendicular to z_i. Thus

$$H(\operatorname{ad}(x) z_i, z_i) = t_{\gamma_i} H(z_i, z_i).$$

Now since $\operatorname{ad}(x)$ is self adjoint and $|\operatorname{ad}(x)| < 1$, it follows that $t_{\gamma_i} < 1$ as desired.

For part (2), let $x = \epsilon_D + y$ with $(y,\lambda_i) = 0$, $i > k$; then

$$y = \sum (t_\gamma a_\gamma + s_\gamma b_\gamma)$$

where $t_\gamma = 0$ if $\gamma = \gamma_{k+1}, \cdots, \gamma_m$. We observe that $[\epsilon_D, z_i] = z_i$ for $i = k+1, \cdots, m$. Then

$$[x, z_i] = z_i + \sum_{\gamma\neq\gamma_i} t_\gamma(\gamma,\gamma_i)/(\gamma_i,\gamma_i)\,(ib_\gamma) \\ + \sum s_\gamma(\gamma,\gamma_i)/(\gamma_i,\gamma_i)\,(-ia_\gamma) + K$$

where K is as above. Thus $\operatorname{ad}(x)z_i = z_i + w$, where we see at once that $H(z_i, w) = 0$. It follows that $|\operatorname{ad}(x)z_i|^2 = |z_i|^2 + |w|^2$ and since $|\operatorname{ad}(x)| < 1$, we must have $w = 0$. This implies that $t_\gamma(\gamma,\gamma_i) = s_\gamma(\gamma,\gamma_i) = 0$ for $i = k+1,\cdots,m$. Then if t_γ or $s_\gamma \neq 0$ we must have $(\gamma,\gamma_i) = 0$ for $i = k+1,\cdots,m$, and so $X_{\pm\gamma} \in \mathfrak{p}(D)$. Thus we must have $y \in \mathfrak{p}_0(D)$.

We recall now the first of Pyatetskiĭ-Shapiro's definition of a boundary component.

Definition. A subset f of $\partial\mathcal{D}$ is an analytic boundary component of $\mathcal{D}$ if it is a regular analytic set in $\mathfrak{p}_0$ contained in $\partial\mathcal{D}$ such that if $\phi(z)$, $|z| < \epsilon$ is a holomorphic curve lying in $\partial\mathcal{D}$ with $\phi(0) \in f$, then $\phi(z) \in f$ for all z.

Hereafter D will denote a set D_k $k = 0, 1, \cdots, m-1$.

THEOREM 6. *If f is any set of the form $k\cdot i_D(s_D)$ there is complex affine subspace H in $\mathfrak{p}_0$ such that $H \cap \bar{\mathcal{D}} = \bar{f}$. The interior of $H \cap \bar{\mathcal{D}}$ in H is f and f is a bounded symmetric domain in H with a Harish-Chandra embedding.*

Proof. We mean by the final statement that we can choose a point (an origin) in H and thereby give H the structure of a complex vector space, and that f is embedded in this vector space by a Harish-Chandra embedding.

Since K acts by unitary transformations, it is enough to prove the theorem for $f = i_D(S_D)$. In this case we let $H = \epsilon_D + \mathfrak{p}_0(D)$ and take ϵ_D as the origin in H. It is clear that $\bar{f} \subset \bar{\mathcal{D}} \cap H$. On the other hand if $P \in \bar{\mathcal{D}} \cap H$, then $|\operatorname{ad}(P)| \leqq 1$. We write $P = \epsilon_D + P_1$ with $P_1 \in \mathfrak{p}_0(D)$. If $x \in \mathfrak{g}(D)$, then $[\epsilon_D, x] = 0$ and so

$$\operatorname{ad}(P)|_{\mathfrak{g}(D)} = \operatorname{ad}(P_1)|_{\mathfrak{g}(D)}.$$

Thus $|\operatorname{ad}(P_1)|_{\mathfrak{g}(D)}| \leqq 1$ and by Lemma 4.5, $P_1 \in z_D(S_D)$. It follows that $P \in \bar{f}$, and that $\bar{f} = \bar{\mathcal{D}} \cap H$. The final statements of the lemma are immediate consequences of Lemma 4.4.

THEOREM 7. *The analytic boundary components of D are exactly the sets f of the form $k\cdot i_D(S_D)$, $k \in K$.*

Proof. We first observe that Theorem 6 implies that any set of the form f above is a regular analytic set since it is locally the zeroes of the hyperplane H. Since the sets of the form f exhaust $\partial\mathcal{D}$ it suffices to prove only that each one of them is an analytic boundary component. Further it is enough just to prove that $i_D(S_D)$ is one.

Let $\phi(z)$, $|z| < \epsilon$ be a holomorphic curve in $\partial\mathfrak{D}$ such that $\phi(0) = p \in f$. Since $G(D)$ maps $\partial\mathfrak{D}$ and f onto themselves and is transitive on f, we may assume without loss of generality that $\phi(0) = \epsilon_D$. Then we can write

$$\phi(z) = \sum(\phi_\beta(z)a_\beta + \psi_\beta(z)b_\beta) \quad (\beta \in P^+)$$

where ϕ_β and ψ_β are harmonic functions. By Lemma 4.6(1) $|\phi_{\gamma_i}(z)| \leqq 1$ since $\phi(z) \in \bar{\mathfrak{D}}$. On the other hand, $\phi_{\gamma_i}(0) = 1$ for $i = k+1, \cdots, m$, and by the maximum principle $\phi_{\gamma_i}(z) = 1$ for all z. Therefore $\phi(z) = \epsilon_D + \phi_1(z)$ where $(\phi_1(z), \lambda_i) = 0$ $i > k$. By Lemma 4.6(2) we must have $\phi_1(z) \in \mathfrak{p}_0(D)$ and so $\phi(z) \in \bar{f}$ in the notation of Theorem 6. We know that f is a bounded domain in H, and if $\phi(z_0) \in \bar{f} - f$ for some z_0, $\phi(z)$ would be entirely contained in the closure of some "boundary component" of f. This is impossible since $\phi(0) = \epsilon_D$. Thus $\phi(z) \in f$ and the theorem is proved.

Pyatetskiĭ-Shapiro has given an alternate definition of boundary components in terms of a metric on $\bar{\mathfrak{D}}$. Let ρ be the Riemannian metric on $\mathfrak{D}$, then if $p, q \in \bar{\mathfrak{D}}$, let

$$\rho(p, q) = \inf(\limsup_n \rho(p_n, q_n))$$

where the infimum runs over all pairs of sequences p_n and q_n which converge to p and q in the Euclidean topology on $\bar{\mathfrak{D}}$. Then ρ is a metric on $\bar{\mathfrak{D}}$. (We allow metric functions to take the value $+\infty$.)

Definition. The mertic components of $\partial\mathfrak{D}$ are the connected components of the mertic ρ on $\partial\mathfrak{D}$.

We shall show that this definition yields the same objects as the previous one. Further we shall give a simple criterion for two points to belong to the same boundary component.

LEMMA 4.7. *Any set of the form $f = k \cdot i_D(S_D)$ is ρ-connected.*

Proof. Since G acts isometrically on $\bar{\mathfrak{D}}$, it is enough to prove that $i_D(S_D)$ is ρ-connected. Let $\mathfrak{g}^{(i)}$ be the three-dimensional complex algebra spanned by X_{γ_i}, $X_{-\gamma_i}$, γ_i, and let $\mathfrak{g}_0^{(i)}$ be its intersection with $\mathfrak{g}_0$. If $\mathfrak{D}' = z(\prod_{i=1}^{m} \exp(g_0^{(i)})K)$ it is known ([8], III) that $\mathfrak{D}'$ is an m-dimensional polycylinder which is embedded totally geodesically in $\mathfrak{D}$. The metric ρ restricted to $\mathfrak{D}'$, ρ' is a Riemannian metric invariant under the connected component $\prod_{i=1}^{m} \exp(g_0^{(i)})$ of analytic automorphisms of $\mathfrak{D}'$. Let ρ' also denote the extension of ρ' to $\bar{\mathfrak{D}}'$. Then it is easy to see that a ρ'-connected subset of

$\partial\mathfrak{D}'$ is ρ-connected as a subset of $\partial\mathfrak{D}$. It is known, however, by explicit calculations that

$$T=\{\sum_{i=1}^{k}r_i\lambda_i+\epsilon_D \quad |r_i|<1\}$$

is a ρ'-connected, set of $\partial\mathfrak{D}'$ containing ϵ_D. Thus T is ρ-connected, and if $k\in K(D)$, $k\cdot T$ is also ρ-connected and contains ϵ_D. Finally it follows that $i_D(S_D)=\bigcup k\cdot T$ $(k\in K(D))$ is ρ-connected.

We recall that g is the homeomorphism of $M=\overline{\phi_B(G/K)}$ onto $\bar{\mathfrak{D}}$. Let $m=g^{-1}$ so that $m(p)$ for $p\in\bar{\mathfrak{D}}$ is the "Poisson measure" on the Shilov boundary S of $\mathfrak{D}$ which represents p. If μ and ν are two mutually absolutely continuous probability measures on S, let

$$\lambda(\mu,\nu)=\operatorname{ess\,sup}(\log(d\nu/d\mu)).$$

LEMMA 4.8. *There is a positive constant c such that for any $p,q\in D$*

$$\lambda(m(p),m(q))\leqq c\rho(p,q).$$

Proof. Since $\lambda(h\cdot\mu,h\cdot\nu)=\lambda(\mu,\nu)$ for $h\in G$ and since ρ is invariant under G, it is enough to prove the lemma when $p=z(K)$ and $q=z(\exp(H)K)$ with $H\in\bar{\Lambda}$. Then $m(p)=\mu$, the K invariant measure on $S\cong G/B(E)$, and $m(q)=a\cdot\mu$ where $a=\exp(H)$.

Let μ_0 is the K invariant measure on G/B, the maximal boundary of Furstenberg, and let h be the projection of G/B onto S. Then $h(\mu_0)=\mu$ and $h(\nu_0)=\nu$ where $\nu_0=a\cdot\mu_0$. It is clear that

$$\sup(d\nu/d\mu)\leqq\sup(d\nu_0/d\mu_0)$$

Now since $a\in\exp(\bar{\Lambda})$, one can see that the maximum value of $d\nu_0/d\mu_0$ occurs at B in G/B and that this maximum value is $\exp(2\alpha(H))$ where α is half the sum of the positive roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$. (cf. [16], p. 281, [3], p. 241). Thus $\lambda(\mu,\nu)\leqq 2\alpha(H)$.

On the other hand, the Riemannian distance in G/K from K to $a\cdot K$ is clearly $\sqrt{(H,H)}$ where $(\ ,\ )$ is the Killing form of $\mathfrak{g}_0$. In order to complete the proof, we choose c so that $\sqrt{(H,H)}\geqq c\,2\alpha(H)$ for $H\in\bar{\Lambda}$.

We can now prove our final result.

THEOREM 8. *The sets of the form $f=k\cdot i_D(S_D)$ are the metric components of $\partial\mathfrak{D}$. Moreover if $p,q\in\mathfrak{D}$, then p and q belong to the same metric component if and only if $m(p)$ and $m(q)$ have the same support.*

Proof. It follows from [13] that $m(p)$ and $m(q)$ have the same support if and only if they lie in the same set $k\cdot i_D(S_D)$. Thus the first and second statements are equivalent. From Lemma 4.7 we know that if $m(p)$ and $m(q)$ have the same support, they lie in the same metric component of $\partial\mathfrak{D}$ since we proved there that $k\cdot i_D(S_D)$ is connected.

It suffices then to prove that if p and q lie in the same metric component, then $m(p)$ and $m(q)$ have the same support. If p and q lie in the same component, then clearly $\rho(p,q)<\infty$. Therefore we can find sequences p_n and q_n in $\mathfrak{D}$ tending respectively to p and q such that $\rho(p_n,q_n)\leqq M<\infty$. By the previous lemma,

$$dm(p_n)/dm(q_n)\leqq \exp(cM)=K,$$

or alternately $Km(q_n)\geqq m(p_n)$. Since $m(q_n)$ and $m(p_n)$ tend to $m(q)$ and $m(p)$ in the weak-* topology, it follows that $Km(q)\geqq m(p)$. By symmetry $Km(p)\geqq m(q)$ and it follows that $m(p)$ and $m(q)$ are mutually absolutely continuous and in particular that they have the same support.

Let $\mathfrak{A}$ denote the Banach algebra of continuous complex valued functions on $\bar{\mathfrak{D}}$ which are holomorphic in $\bar{\mathfrak{D}}$. The maximal ideal space of $\mathfrak{A}$ is $\bar{\mathfrak{D}}$. In [17] Gleason has defined the notion of a part of a maximal ideal space. If p and q are in $\bar{\mathfrak{D}}$, then we say that $p\sim q$ if

$$\sup_{|f|=1,\,f\in\mathfrak{A}}|f(p)-f(q)|<2,\qquad |f|=\sup_{x\in\bar{\mathfrak{D}}}|f(x)|.$$

It is proved that $\sim$ is an equivalence relation and the equivalence classes of this relation are defined to be the Gleason parts of $\bar{\mathfrak{D}}$. The following is not unexpected and not difficult to prove.

THEOREM 9. *The Gleason parts of $\bar{\mathfrak{D}}$ are $\bar{\mathfrak{D}}$ together with the boundary components of $\partial\mathfrak{D}$.*

Appendix.

Koranyi and Wolf [11] say that an irreducible Cartan domain G/K is of tube type if it can be realized as a tube domain over a self dual cone. They show that G/K is a tube domain if and only if alternative (1) of Theorem 2 holds, or equivalently, if and only if $-\frac{1}{2}\gamma_1$ is not in $\pi(R)$ (R is the set of roots of $\mathfrak{g}$ with respect to $\mathfrak{a}$). In the following, boundary component means only those of positive dimension.

THEOREM. (1) *If G/K is of tube type, then each boundary component is of tube type. If G/K is not of tube type then no boundary component is of tube type.*

(2) *All boundary components are classical domains.*

Proof. For our purposes, we may take a boundary component f of the form $i_D(S_D)$ with D non void. Then $D = D_k$ with $k > 0$ and the connected group of isometries of S_D is essentially $G(D)$. By Lemma 4.3 and the discussion preceding it, the roots of $\mathfrak{g}(D)$ with respect to $\mathfrak{a}(D)$ are those roots β of $\mathfrak{g}$ with respect to $\mathfrak{a}$ such that $(\beta, \gamma_i) = 0$ for $i = k+1, \cdots, m$.

(1) If G/K is of tube type we have alternative (1) in Theorem 2 for $\pi(R)$, and by the remark above, alternate (1) holds also for the roots of $\mathfrak{g}(D)$. Thus S_D is of tube type. The second statement of part (1) follows in the same way.

(2) We consider the case when G/K is exceptional. There are two exceptional domains: one of dimensional 16 which is not of tube type, and one of dimension 27 which is of tube type. If G/K is exceptional of dimension 16, all boundary components are of lower dimension and hence classsical. If G/K is exceptional of dimension 27, all boundary components are of lower dimension and of tube type, hence classical.

To complete the proof, we must show that if G/K is classical, then each $G(D)/K(D)$ is also classical. We recall that $\alpha_1, \cdots, \alpha_n$ is a fundamental system of roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$ using the ordering described immediately after Theorem 2′. Then some subset S of $\alpha_1, \cdots, \alpha_n$ will be a fundamental system of roots of $\tilde{\mathfrak{g}}(D)$ with respect to $\mathfrak{h}(D)$ (cf. [15]). Let $\tilde{\mathfrak{g}}(D)$ be the unique non-compact component of $\mathfrak{g}(D)$. Then if $G(D)/K(D)$ is not classical, $\tilde{\mathfrak{g}}(D)$ is an exceptional Lie algebra. If β is its highest root, then $\beta = \sum n_i \alpha_i$ $(\alpha_i \in S)$. It is known that for at least one i, $n_i > 2$. On the other hand, β is a root of $\mathfrak{g}$, and since $\mathfrak{g}$ is classical, all $n_i \leqq 2$. This contradiction establishes the Theorem.

UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

REFERENCES.

[1] E. Cartan, "Sur les domaines bornés homogènes de l'espace de n variables complexes," *Abhandlungen aus dem Mathematische Seminar der Universität Hamburg*, vol. 11 (1935), pp. 116-162.

[2] H. Furstenberg, "A Poisson formula for semi-simple Lie groups," *Annals of Mathematics*, vol. 77 (1963), pp. 335-386.

[3] Harish-Chandra, "Representations of semi-simple Lie groups I," *Transactions of the American Mathematical Society*, vol. 75 (1952), pp. 185-243.

[4] ———, "Representations of semi-simple Lie groups VI," *American Journal of Mathematics*, vol. 78 (1956), pp. 564-628.

[5] S. Helgason, *Differential Geometry and Symmetric Spaces*, Academic Press, New York (1962).

[6] L. K. Hua, *Harmonic Analyses of several complex variables in the classical domains*. Mimeographed notes.

[7] R. Hermann, "A Poisson kernel for certain homogeneous spaces," *Proceedings of the American Mathematical Society*, vol. 12 (1961), pp. 892-899.

[8] ———, "Geometric aspects of potential theory in bounded symmetric domains, I," *Mathematische Annalen*, vol. 148 (1962), pp. 349-366. II, III *to appear.*

[9] N. Jacobson, *Lie Algebras*. Interscience, New York, 1962.

[10] M. Koecher, *Jordan Algebras and their Applications*, Mimeographed notes, University of Minnesota, 1962.

[11] A. Koranyi and J. Wolf, "Realization of Hermitian Symmetric Spaces as generalized half planes." (*forthcoming*)

[12] D. Lowdenslager, "Potential theory in bounded symmetric homogeneous complex domains," *Annals of Mathematics*, vol. 67 (1958), pp. 467-484.

[13] C. Moore, "Compactifications of Symmetric Spaces I." (*forthcoming*)

[14] I. I. Pyatetskiĭ-Shapiro, *Geometry of Classical Domains and the theory of Automorphic functions*, Moskow, 1961.

[15] I. Satake, "On Representations and Compactifications of Symmetric Spaces," *Annals of Mathematics*, vol. 71 (1960), pp. 77-110.

[16] Harish-Chandra, "Spherical functions on a semi-simple Lie group," *American Journal of Mathematics*, vol. 80 (1958), pp. 241-310.

[17] A. Gleason, *Function Algebras, Seminars on Analytic Functions*, Princeton, 1957.

SHEAVES DEFINED BY DIFFERENTIAL EQUATIONS AND APPLICATION TO DEFORMATION THEORY OF PSEUDO-GROUP STRUCTURES.

By MASATAKE KURANISHI.

For a vector bundle E over a manifold V, denote by $S(E)$ the sheaf of germs of local cross-sections of E. By using local coordinate we may consider linear differential equations on E. When $V = \boldsymbol{R}^n$ and $E = V \times \boldsymbol{R}^m$, the linear differential equations on E are linear differential equations on m unknown functions in n variables. The germs of solutions of a linear differential equation on E forms a subsheaf, say G, of $S(E)$. The importance of such subsheaves G can be seen, for instance, in the recent works of Kodaira and Spencer in the deformation theory of pseudo-group structures. In the first half of the present paper, we study some of properties of such sheaves. Since it is very difficult to treat the general case, we restrict ourselves to the category of real analyticity and to the case where inconvenient degeneracy does not occur in the differential equations. In our argument these restrictions are necessary because in the last analysis we depend on Cartan's theorem of solvability of involutive real analytic differential equations. G together with its defining linear differential equation which is nondegenerate, is called a (S)-sheaf. Our main result is the existence of quotient (S)-sheaves of (S)-sheaves by (S)-subsheaves. In the second half the writer gives concise description of how the above results can be applied to deformation theory of transitive continuous pseudo-group compact structures. Since our argument follows the more or less well known line developed in [3], details are omitted. It is noted that another approach, which is still within the same framework as ours but which is more differential geometric, is published in [5] by D. C. Spencer.

In the present paper every notions are assumed to be in the category of real analyticity. So when we say manifolds or mappings for example, we mean always manifolds or mappings which are real analytic.

1. Differential equations and sheaves of type (S). Let N and N' be manifolds. Take a mapping π of N' to N. We say that the triple (N', N, π) is a fibered manifold when π is surjective and the rank of $d\pi$ at each point of N' is equal to the dimension of N. When that is the case, for any point p'

Received October 4, 1963.

of N' there is a neighborhood U' of p' and a coordinate $(x^1,\cdots,x^n,y^1,\cdots,y^m)$ defined on U' such that there is a coordinate $('x^1,\cdots,'x^n)$ defined on $\pi(U')$ satisfying the equality: $'x^i = x^i \circ \pi$. Such a coordinate (x,y) is called a coordinate of the fibered manifold (N',N,π). As far as no confusion can occur, we use x^i instead of $'x^i$. Let $J^l(N',N,\pi)$ be the space of l-jets of local cross-sections of (N',N,π). $J^l(N',N,\pi)$ has the canonical manifold structure. Denote by α and β the source and the target mappings, respectively. Setting $J^l = J^l(N',N,\pi)$, we see easily that

$$(J^l,N,\alpha),\ (J^l,N',\beta),\ \text{and}\ (J^l,N\times N',\alpha\times\beta)$$

are fibered manifolds. We have the canonical projection $\rho^l{}_{l'}$ of J^l onto $J^{l'}$ $(l\geqq l')$. $(J^l,J^{l'},\rho^l{}_{l'})$ is also a fibered manifold. J^0 can be canonically identified with N, and α with π. Let (x,y) be a coordinate defined on U' of (N',N,π). Then J^l has the canonical coordinate (x,y,p) defined on $\beta^{-1}(U')$ as follows: Take $X = j_z{}^l(f)$ in $\beta^{-1}(U')$, where f is a cross-section over a neighborhood of z. Express f in terms of coordinate (x,y) as $y^\lambda = f^\lambda(x)$, $\lambda = 1,\cdots,m$. Then

$$\begin{aligned}&p^\lambda{}_{i_1\cdots i_\nu}(X) = (\delta f^\lambda/\delta x^{i_1}\cdots\delta x^{i_\nu})_{x=x(z)}\\&x^i(X) = x^i(z),\qquad y^\lambda(X) = y^\lambda(f(z)),\\&i,i_1,\cdots,i_\nu = 1,\cdots,n\ \text{and}\ \nu\leqq l.\end{aligned}$$

Thus $p^\lambda{}_{i_1\cdots i_\nu}$ is symmetric in $i_1,\cdots,i_\nu$. The coordinate (x,y,p) is said to be associated with (x,y). For a cross-section f of (N',N,π) over an open set U of N, denote by $j^l(f)$ the cross-section: $U\ni z\to j_z{}^l(f)\in J^l$ of (J^l,N,α) over U. Let $\mathcal{U}$ be an open subset of J^l. Denote by $\pi(\mathcal{U})$ the space of Pfaffian forms w on $\mathcal{U}$ such that $j^l(f)^*w = 0$ for any local cross-section f of (N',N,π) with the image of $j^l(f)$ in $\mathcal{U}$. If $\mathcal{U}\supset\mathcal{U}'$, the restriction mapping send $\pi(\mathcal{U})$ to $\pi(\mathcal{U}')$. Thus the system $\Pi(\mathcal{U})$ defines a subsheaf, say $\Pi(1,(N',N,\pi))$, of the sheaf ΛJ^l of germs of differential forms on J^l. The following propositions are well known.

PROPOSITION 1.1. *$\Pi(l,(N',N,\pi))$ is locally finitely generated over the sheaf $\Lambda^0 J^l$ of germs of functions. More precisely, if (x,y) is a coordinate of (N',N,π) defined on U' and if $\mathcal{U}'$ is an open subset of $\beta^{-1}(U')$), then $\Pi(\mathcal{U}')$ is generated by*

$$\begin{aligned}&w^\lambda = dy^\lambda - p^\lambda{}_j\,dx^j\\&w^\lambda{}_{j_1\cdots j_\nu} = dp^\lambda{}_{j_1\cdots j_\nu} - p^\lambda{}_{j_1\cdots j_\nu j}\,dx^j\quad(\nu\leqq l-1).\end{aligned}$$

PROPOSITION 1.2. *Let F be a local cross-section of (J^l, N, α). Then there is a local cross-section f of (N', N, π) such that $F = j^l(f)$ if and only if $F^* w = 0$ for any open set $\mathcal{U}$ of J^l and for any w in $\pi(\mathcal{U})$.*

Definition 1.1. By a partial differential equation of order l on (N', N, π), we mean an open subset $\mathcal{U}$ of $J^l(N', N, \pi)$ and a subsheaf Σ of ideals of $\Lambda^0\mathcal{U}$ which is locally finitely generated over $\Lambda^0\mathcal{U}$. A local cross-section f of (N', N, π) is called an integral of Σ when $j^l(f)$ is in the set of zeros of Σ.

By a locally closed subvariety W of a manifold N'' we mean a subset W of N'' satisfying the following condition: For every point p of W we can find an open neighborhood $U(p)$ of p in N'' such that $W \cap U(p)$ is the set of zeros of a finite number of functions defined on $U(p)$. When this is the case, let U be the union of all such $U(p)$ where p runs through W. Then W is a closed subset of U. Thus for a locally closed subvariety W of N'' there is an open subset, say U, of N'' such that W is a closed subset of U. Let A be a locally closed subvariety of J^l. Take an open subset $\mathcal{U}$ in J^l such that A is closed in $\mathcal{U}$. Denote by $\Sigma(A)$ the subsheaf of $\Lambda^0\mathcal{U}$ consisting of all germs of local functions which vanish on A. $\Sigma(A)$ is a subsheaf of ideals of $\Lambda^0\mathcal{U}$. Then $\Sigma(A)$ is a differential equation of order l on (N', N, π), which is said to be associated with A. A will be also referred as a differential equation, meaning $\Sigma(A)$. $\Sigma(A)$ is not uniquely determined by A because the above choice of $\mathcal{U}$ is not unique. However, this indetermination is of no importance to our discussion. For a differential equation Σ of order l on (N', N, π), the set $I^0\Sigma$ of zeros of Σ is a locally closed subvariety of J^l. Clearly $\Sigma(I^0\Sigma) \supseteq \Sigma$.

Let (x, y) be a coordinate in (N', N, π) defined on U'. Let g be a function defined on an open subset $\mathcal{U}$ of $\beta^{-1}(U') \subseteq J^l$. Then $g \circ \rho^{l+1}{}_l$ is a function on $(\rho^{l+1}{}_l)^{-1}(\mathcal{U})$ and $d(g \circ \rho^{l+1}{}_l)$ is a linear combination of $dx^i, dy^\lambda, dp^\lambda{}_{j_1 \cdots j_\nu}$ $(\nu \leqq l)$. Hence by Prop. 1.1,

$$d(g \circ \rho^{l+1}{}_l) \equiv g_j \, dx^j \qquad (\operatorname{mod} \Pi(l+1; (N', N, \pi))) \tag{1}$$

where g_j is a function defined on $(\rho^{l+1}{}_l)^{-1}(\mathcal{U})$. We set

$$\delta^{\#}{}_x{}^j(g) = g_j. \tag{2}$$

When (x, y) is changed to (z, w), $dx^j = a_k{}^j dz^k$ and so

$$\delta^{\#}{}_z{}^j(g) = a_j{}^k(\delta^{\#}{}_x{}^k(g)).$$

Let F be an ideal of the ring $R(\mathcal{U})$ of functions on $\mathcal{U} \subseteq J^l$. Denote by $P(F)$ the ideal in $R(\rho^{l+1}{}_l)^{-1}(\mathcal{U}))$ generated by $g \circ \rho^{l+1}{}_l$ and $\delta^{\#}{}_x{}^j(g)$, $j = 1, \cdots, n$, where g runs through F. The above rule for changes of coordinates shows that $P(F)$ is independent of the coordinate used in the construction. Let Σ be a subsheaf of ideals of $\Lambda^0\mathcal{U}$. For an open subset $\mathcal{U}'$ of $\mathcal{U}$ the space $\Gamma(\mathcal{U}',\Sigma)$ of sections of Σ over $\mathcal{U}'$ is an ideal of $R(\mathcal{U}')$, and so $P(\Gamma(\mathcal{U}',\Sigma))$ is defined. If $\mathcal{U}'\supseteq\mathcal{U}''$, the restriction mapping sends $P(\Gamma(\mathcal{U}',\Sigma))$ to $P(\Gamma(\mathcal{U}'',\Sigma))$. Thus the system $\{P(\Gamma(\mathcal{U}',\Sigma))\}$ defines a subsheaf of ideals of $\Lambda^0((\rho^{l+1}{}_l)^{-1}(\mathcal{U}))$, which will be called the standard prolongation of Σ and will be denoted by $P(\Sigma)$. $P(\Sigma)$ is a differential equation of order $l+1$ on (N',N,π). Let A be a locally closed subvariety of J^l. If $P(\Sigma(A))=\Sigma(I^0P(\Sigma(A)))$, we say that the standard prolongation of A can be defined. When this is the case, $I^0P(\Sigma(A))$ is denoted by $P(A)$. The following proposition and lemma are easy to check.

PROPOSITION 1.3. *Let Σ be a differential equation on (N',N,π). A local cross-section f of (N',N,π) is a solution of Σ if and only if f is a solution of $P(\Sigma)$.*

LEMMA 1.1. *Take a coordinate (x,y) defined on U' of (N',N,π). Let f be a local cross-section of (N',N,π) such that the image of f is in U'. Denote by ϕ a function defined on a neighborhood of the image of $j^l(f)$. Then*

$$(\delta/\delta x^k)\phi(j^l(f))=[\delta^*_{x^k}\phi](j^{l+1}(f)).$$

Let Σ be a differential equation of order l on (N',N,π). Denote by $\mathcal{U}$ the open subset of J^l on which Σ is defined. Let Σ^* be the subsheaf of ideals (closed under the exterior differentiation) of $\Lambda\mathcal{U}$ generated by Σ and by the restriction to $\mathcal{U}$ of $\Pi(l,(N',N,\pi))$. We say that Σ is in involution, when Σ^* is an exterior differential system in involution. It is well known that if Σ is in involution then so is $P(\Sigma)$. A locally closed subvariety A is said to be in involution when $\Sigma(A)$ is in involution. If this is the case, A is a locally closed submanifold, $P(A)$ is defined, non-empty (provided A is non-empty), and is in involution. Hence $P(A)$ is also a locally closed submanifold.

Definition 1.2. Let A^l be a non-empty locally closed submanifold of $J^l(N',N,\pi)$, $l\geqq l_1$. Take X^l in A^l. We say that the sequence $\{A^l,X^l;l\geqq l_1\}$ is normal when the following conditions are satisfied for each $l\geqq l_1$:

(1) $\rho^{l+1}{}_l(X^{l+1})=X^l$,

(2) there is an open neighborhood $\mathcal{U}^{l+1}$ of X^{l+1} such that the restriction of $P(\Sigma(A^l))$ to $\mathcal{U}^{l+1}$ is contained in $\Sigma(A^{l+1})$,

(3) we can find an open neighborhood $'\mathcal{U}^l$ of X^l such that $(\mathcal{U}^{l+1}\cap A^{l+1}, {}'\mathcal{U}^l\cap A^l, \rho^{l+1}{}_l)$ is a fibered manifold,

(4) there is an open neighborhood $\mathcal{U}$ of X^{l_1} such that $(\mathcal{U} \cap A^{l_1}, \alpha(\mathcal{U}), \alpha)$ is a fibered manifold.

A normal sequence is called strictly normal when $P(A^l)$ is defined and coincides with A^{l+1} on a neighborhood of X^{l+1} for any $l \geqq l_1$. The following is proved in [4].

THEOREM 1. *Let* $\{A^l, X^l; l \geqq l_1\}$ *be a normal sequence. Then there is an integer* $l_2 \geqq l_1$ *such that, for any* $l \geqq l_2$, *there is a neighborhood* $\mathcal{U}^l$ *of* X^l *such that* $\mathcal{U}^l \cap A^l$ *is in involution and* $P(\Sigma(A^l))$ *coincides with* $\Sigma(A^{l+1})$ *on a neighborhood of* X^{l+1}.

By the existence theorem of solutions for involutive system, we have the following corollary:

COROLLARY. *Under the same assumptions as in Theorem* 1, *for any fixed* $l \geqq l_1$ *and for any* X *in* A^l *sufficiently near* X^l, *there is a solution* f *of the differential equation* $\Sigma(A^l)$ *such that* $j_p{}^l(f) = X$ *where* $p = \alpha(X)$.

Let (E, N, π) be a vector bundle. Then $(J^l(E, N, \pi), N, \alpha)$ is also a vector bundle over N. As far as no confusion can occur, we write $J^l(E)$ instead of $J^l(E, N, \pi)$. By a linear differential equation of order l in (E, N, π) we mean a vector sub-bundle (B, N, α) of $(J^l(E), N, \alpha)$. Denote by $O_p{}^l$ $(p \in N)$ the l-jet at p of the zero-section. A sequence B^l $(l \geqq l_1)$ of linear differential equation of order l in (E, N, π) is said to be (strictly) normal when the sequence $\{B^l, O_p{}^l; l \geqq l_1\}$ is (strictly) normal for any p in N. A linear differential equation B of order l_1 in (E, N, π) is called weakly involutive when there is a strictly normal sequence $\{B^l; l \geqq l_1\}$ such that $B = B^{l_1}$. Let $\mathfrak{S}(E)$ be the sheaf of germs of local cross-sections of (E, N, π). Denote by $\mathfrak{S}(E, B)$ the subsheaf (of $\mathfrak{S}(E)$) of germs of integrals of B. If B is weakly involutive, $\mathfrak{S}(E, B) = \mathfrak{S}(E, B^l)$ for $l \geqq l_1$ by Prop. 1.3. Let (E', N', π') be a vector bundle.

Definition 1.3. Let B (resp. B') be a linear differential equation in E (resp. E'). Denote by f a mapping of N to N'. A continuous mapping a of $\mathfrak{S}(E, B)$ to $\mathfrak{S}(E', B')$ is said to be a linear differential mapping (over f) when we can find integers l_0 and k such that for any $l \geqq \max(l_0, k)$ there is a fiber preserving mapping $a_l{}^*$ of $J^l(E)$ into $J^{l-k}(E')$, which induces f on the base spaces and linear mappings on fibers, and which satisfies the following condition: For any local cross-section f of $\mathfrak{S}(E, B)$

$$j^{l-k}(a(f)) = a_l{}^*(j^l(f)).$$

LEMMA 1.2. *Keeping the notations in Def.* 1.3, *assume further that* B' *is weakly involutive and that* a *is a surjective mapping of* $\otimes(E,B)$ *to* $\otimes(E',B')$. *Let* l_1 (*resp.* l_1') *be the order of* B (*resp.* B'). *Then for sufficiently large* l *the image of* $P^{l-l_1}(B)$ *by* a_l^* *is equal to* $P^{l-l_1'-k}(B')$, *where* P^k *denotes* k *times iteration of the operation* P.

Proof. By Prop. 1.3 and Theorem 1, we can assume that B' is in involution. Otherwise replace B' *by* $P^r(B')$ with a sufficiently large r. Then our lemma is an immediate corollary of the existence theorem of solutions for involutive systems.

Definition 1.4. Keeping the notation in Def. 1.3, assume further that B is weakly involutive. Then we say that a is regular when, for a suitable choice of a_l^* (i.e., of k), $C^l = B^l \cap$ (the kernel of a_l^*) is a vector sub-bundle of $J^l(E)$ and $(C^{l+1}, C^l, \rho^{l+1}{}_l)$ is a fibered manifold for sufficiently large l, where $B^l = P^{l-l_1}(B)$ and l_1 is the order of B.

Definition 1.5. By a (S)-structure $\mathcal{F}$ on a sheaf F (of modules) over N we mean the following: (1) For any open subset U of N with compact closure we are given a vector bundle $(E(\mathcal{F},U),U,\pi)$ and a weakly involutive linear differential equation $B(\mathcal{F},U)$ in $(E(\mathcal{F},U),U,\pi)$ together with a homeomorphism $\iota_{\mathcal{F}U}$ of F over U onto $\otimes(E(\mathcal{F},U),B(\mathcal{F},U))$. (2) for any $U \supset U'$ the mapping $\iota_{\mathcal{F}U} \circ \iota_{\mathcal{F}U'}^{-1}$ is a regular linear differential mapping over the canonical injection of U' into U.

Let $\mathcal{F}$ be a (S)-structure on F. We regard $\mathcal{F}'$ is equal to $\mathcal{F}$ when $P^k(B(\mathcal{F},U)) = P^{k'}(B(\mathcal{F}',U))$ for a suitable choice of k and k'.

F will be called the underlying sheaf of the (S)-structure $\mathcal{F}$. When there is no possibility of confusion, we use F to denote the (S)-structure as well as the underlying sheaf. A sheaf with a (S)-structure will be referred as a (S)-sheaf.

Remark. When N is compact, we may restrict ourselves to a special case of (S)-structure where $E(\mathcal{F},U)$ (resp. $B(\mathcal{F},U)$ is the restriction to U of a vector bundle E over N (resp. a weakly involutive differential equation B). Then the subsequent arguments are somewhat, but not essentially, simplified.

Definition 1.6. Let F_1 and F_2 be (S)-sheaves over N. Let a be a continuous mapping of F_1 to F_2. We say that a is a (S)-mapping when the following is satisfied: For any open subset U of N with compact closure $\iota_{F^2U} \circ a \circ \iota_{F^1U}^{-1}$ is a regular linear differential mapping of $\otimes(E(F_1,U),B(F_1,U))$ into $\otimes(E(F_2,U))$.

Let F_1 be a subsheaf of a sheaf F with (S)-structure. We say that F_1 is a (S)-subsheaf of the (S)-structure when there is a weakly involutive sub-bundle $B_1(U)$ of $B(F, U)$ for any open subset U of N with compact closure such that ι_{FU} induces a homomorphism of F_1 over U onto $\Theta(E(F, U), B_1(U))$, provided we replace $B(F, U)$ by its standard prolongation if necessary. Then by Lemma 1.2, we see that by associating $E(F, U), B_1(U)$, and the restriction of ι_{FU} to F_1 over U for each U we have a (S)-structure on F_1. In the following, a (S)-subsheaf of a (S)-sheaf will be always considered as a (S)-sheaf in the way mentioned above. It is clear that the canonical injection of F_1 into F is a (S)-mapping. Let G_i $(i = 1, 2, 3)$ be a (S)-sheaf over N. Denote by b (resp. c) a (S)-mapping of G_1 to G_2 (resp. G_2 to G_3). Assume that b is surjective. Then it is easy to check that the composition $c \circ b$ is again a (S)-mapping. Let E be a vector bundle over N. Then the sheaf $\Theta(E)$ has an obvious (S)-structure. Then the canonical mapping j^l of $\Theta(E)$ into $\Theta(J^l(E))$ is a (S)-mapping. Take a homomorphism a of a vector bundle E to a vector bundle E' such that the kernel is a vector sub-bundle of E and which induces the identity mapping on the base space. Then a induces a mapping of $\Theta(E)$ to $\Theta(E')$. It is not hard to check that this mapping is a (S)-mapping.

2. Properties of (S)-sheaves.

THEOREM 2. *Let F_1 and F_2 be (S)-sheaves over a manifold N. Take a continuous mapping a of F_1 to F_2. If a is a (S)-mapping, then the kernel of a and the image of a are (S)-subsheaves of F_1 and F_2, respectively.*

Proof. Take an open subset U of N with compact closure. We can assume that F_i over U is equal to $\Theta(E_i, B_i)$, where E_i is a vector bundle over U and B_i is a weakly involutive differential equation in E_i $(i = 1, 2)$. Moreover, we can assume that a over U is a regular linear differential mapping. Using the notations in Definition 1.4, we set $C^l = B_1{}^l \cap$ (the kernel of $a_l{}^*$). By the definition, C^l is a linear differential equation in E_1 and the kernel of a over U is equal to $\Theta(E_1, C^l)$. Hence, in order to prove that the kernel of a is a (S)-subsheaf it is sufficient to show that, for any subset U' of U such that the closure of U' is in U, C^l over U' is weakly involutive for sufficiently large l. (Note that the closure of U is contained in an open set U_1 with compact closure). Because of Theorem 1, to prove the above, it is sufficient to show that $P(\Sigma(C^l)) \subseteq \Sigma(C^{l+1})$. To check this, take X in C^{l+1} and a function ϕ defined on a neighborhood of $X' = \rho^{l+1}{}_l(X)$ in $J^l(E_1)$ and which vanishes in the intersection with C^l. Let (x, w) be a vector bundle coordinate of $J^{l-k}(E_2)$. We have to show that $[\delta^{\#}{}_{x^i}\phi](X) = 0$. By the definition of

C^l, ϕ is a linear combination of $w^\lambda \circ a_l^*$. Therefore we can assume that $\phi = w^\lambda \circ a_l^*$. Take a local cross-section f of E_1 such that $X = j_p^{l+1}(f)$. Then by Lemma 1.1, we have

$$\begin{aligned}[\delta^{\#}_{x^i}\phi](X) &= (\delta/\delta x^i)(w^\lambda(a_l^*(j^l(f)))) \\ &= (\delta/\delta x^i)(w^\lambda(j^{l-k}(a(f))))\end{aligned}$$

which is one of fiber component of a vector bundle coordinate of $j_p^{l+1-k}(a(f)) = a_{l+1}^*(X)$. On the other hand, since X is in C^{l+1}, $a_{l+1}^*(X) = 0$ and hence it follows that $[\delta^{\#}_{x^i}\phi](X) = 0$. Thus $P(\Sigma(C^l)) \subseteq \Sigma(C^{l+1})$, and so we proved that the kernel of a is a (S)-subsheaf.

Set $A^l = a_{l+k}^*(B_1^{l+k})$. By the condition in Definition 1.4, A^l is a vector subbundle of $J^l(E_2)$. It is clear that the image of $\Theta(E_1, B_1)$ by a is in $\Theta(E_2, A^l)$. We claim that $P(\Sigma(A^l)) \subseteq \Sigma(A^{l+1})$. Namely, take $X = a_{l+k+1}^*(Y)$, $Y \in B_1^{l+1+k}$, in A^{l+1}. Denote by ϕ a function defined on a neighborhood of $X' = \rho^{l+1}{}_l(X)$ in $J^l(E_2)$ and which vanishes on the intersection with A^l. Let (x) be a coordinate in N on a neighborhood of $x_0 = \alpha(X)$. By the existence theorem of solutions of involutive system, we can find a cross-section f of $\Theta(E_1, B_1)$ such that $j_{x_0}^{l+1+k}(f) = Y$. Then $X = j_{x_0}^{l+1}(a(f))$ and $0 = \phi(a_{l+k}^*(j^{l+k}(f)) = \phi(j^l(a(f)))$. Hence

$$[\delta^{\#}_{x^i}\phi](X) = (\delta(\phi(j^l(a(f))))/\delta x^i)_{x=x_0} = 0.$$

This shows that $P(\Sigma(A^l)) \subseteq \Sigma(A^{l+1})$. Therefore by Theorem 1.1 there is an integer l_3 such that for $l \geqq l_3$ A^l is weakly involutive and $\Theta(E_2, A^l)$ is equal to $\Theta(E_2, A^{l_3})$. Thus, in order to prove that the image of F_1 by a is a (S)-sheaf, it is sufficient to show that the image of $\Theta(E_1, B_1)$ is equal to $\Theta(E_2, A^{l_3})$.

Now, take a local cross-section g of $\Theta(E_2, A^{l_3})$. Set

$$K^l = \{X \in B_1^l ; a_l^*(X) = j_p^{l-k}(g), p = \alpha(X)\},$$

where p runs through the domain of g. K^l is a locally closed subvariety of $J^l(E_1)$. Take X_0 in K^l with $p = \alpha(X_0)$. Then clearly $\alpha^{-1}(p) \cap K^l = X_0 + C^l$. Hence K^l is a locally closed submanifold of $J^l(E_1)$ and $(K^{l+1}, K^l, \rho^{l+1}{}_l)$ is a fibered manifold. Set $Y^l = j_p{}^l(g)$ and choose X^l such that $a_l^*(X^l) = Y^{l-k}$ and $\rho^{l+1}{}_l(X^{l+1}) = X^l$. If we can prove that for a sufficiently large l there is a local cross-section f of $\Theta(E_1, K^l)$ such that $j_p{}^l(f) = X^l$, then f is a cross-section of $\Theta(E_1, B_1)$ and $a(f) = g$ on a neighborhood of p and so we can confirm that the image of a is a (S)-subsheaf. To see that such f exists, again by Theorem 1.1 it is sufficient to show that $P(\Sigma(K^l)) \subseteq \Sigma(K^{l+1})$. Take a coordinate (x, w) of $J^{l-k}(E_2)$ and a function ϕ on an open subset in $J^l(E_1)$ which vanishes on the intersection with K^l. Then ϕ is a linear combination of functions ψ^λ, where $\psi^\lambda(W) = w^\lambda(a_l{}^*(W)) - w^\lambda(j_{\alpha(W)}{}^{l-k}(g))$. Then the similar argument used above works here. Thus we proved that the image by a is a (S)-subsheaf.

THEOREM 3. *Let F be a (S)-sheaf. Take a (S)-subsheaf G of F. Then the quotient sheaf F/G has a (S)-structure such that the canonical projection of F to F/G is a (S)-mapping. Moreover such (S)-structure on F/G is unique up to isomorphism.*

Proof. We can assume without loss of generality that $F = \mathfrak{S}(E, A)$ and $G = \mathfrak{S}(E, B)$, where E is a vector bundle and A, B are weakly involutive linear differential equation of order l on E. In order to prove the existence of a (S)-structure on G/F satisfying the required condition, we set $B^k = P^{k-l}(B)$ and $A^k = P^{k-l}(A)$ for $k \geqq l$.

We have (S)-mappings

$$j^l: \mathfrak{S}(E) \to \mathfrak{S}(J^l(E)) \text{ and } b: \mathfrak{S}(J^l(E)) \to \mathfrak{S}(J^l(E)/B^l),$$

where b is the canonical projection. Set $c = b \circ j^l$. Clearly c is a linear differential mapping and represented by a vector bundle mappings $c_k{}^*$ of $J^k(E)$ into $J^{k-l}(J^l(E)/B^l)$ for $k \geqq l$. By direct checking, we see easily that the kernel of $c_k{}^*$ is B^k. Since $B^k \subseteq A^k$ and B^k is weakly involutive, it follows by the definition that $c: F \to \mathfrak{S}(J^l(E)/B)$ is a (S)-mapping, where we denote by c the restriction of the original c to $F = \mathfrak{S}(E, A)$. Since the image of this c is isomorphic to F/G, it follows by Theorem 2 that F/G is a (S)-sheaf.

In order to show the uniqueness, let H_i be a (S)-sheaf together with a surjective (S)-mapping c^i of F to H_i with kernel G, $i = 1$ and 2. We may assume that $H = \mathfrak{S}(E_i, B_i)$ and c^i is induced by a mapping $c_l{}^{i*}$ of $J^l(E)$ to $J^{l-k_i}(E_i)$. Denote by $B_i{}^l$ the prolongation of B_i in $J^l(E_i)$. For a sufficiently large l take an element X in $B_1{}^l$. By Lemma 1.2 and the surjectivity of c^1, there is an element Y in B^{l+k_1} such that $c^1{}_{l+k_1}{}^*(Y) = X$. Set $\tau_l(X) = c^2{}_{l+k_1}{}^*(Y)$. We claim that τ_l is a well defined bundle mapping of $B_1{}^l$ to $B_2{}^{l+k_1-k_2}$. Namely, since the kernel of the mapping $c^1{}_l{}^*$ on B^l is the equation of G as is shown in the proof of Theorem 2, for any W in B^{l+k_1} such that $c^1{}_{l+k_1}{}^*(W) = 0$ there is a cross-section g of G over a neighborhood of $\alpha(W) = p$ such that $W = j_p{}^{l+k_1}(g)$. Hence $c^2{}_{l+k_1}{}^*(W) = j_p{}^{l+k_1-k_2}(c^2(g)) = 0$. Therefore τ_l is well defined. τ_l is surjective because we can go backward using Lemma 1.2. Extending τ_l to a bundle mapping of $J^l(E_1)$ to $J^{l+k_1-k_2}(E_2)$ we obtain a (S)-mapping $c^\#$ of H_1 to H_2 such that $c_2 = c^\# \circ c_1$. This shows that (S)-

structures on F/G which satisfy the required conditions are unique up to isomorphism, q. e. d.

3. Deformations of compact pseudo-group structures. Let Γ be a transitive continuous pseudo-group operating on R^n. Take a Γ-structure V defined on a compact manifold V. Then the structure induces a continuous pseudo-group $\Gamma(V)$ on V, i. e. the set of local transformations of V such that their expressions in terms of Γ-structure coordinates of V belong to Γ. Denote by $\mathcal{L}$ the sheaf of germs of local vector fields belonging to $\Gamma(V)$. Then there is a weakly involutive linear differential equation B of order l_1 in the tangent vector bundle T of V such that $\mathcal{L}$ is equal to $\Theta(T,B)$. For $l \geqq l_1$, denote by B^l the prolongation of B to $J^l(T)$. Then the arguments in § 2 (especially the proof of Theorem 3) show that we have an exact sequence

$$(1) \qquad 0 \to \Theta(T,B) \xrightarrow{i} \Theta(T) \xrightarrow{D} \Theta(J^l(T)/B^l) \xrightarrow{D} \Theta(J^k(J^l(T)/B^l)/C^k)$$

where i is the canonical injection, k is an integer, C^k is a vector sub-bundle of $J^k(J^l(T)/B^l)$, and D's are the composition of the canonical projections and the jet mappings. Moreover, the image by D of $\Theta(T)$ is equal to $\Theta(J^l(T)/B^l, C^k)$. We set

$$E_1 = J^l(T)/B^l, \qquad E_2 = J^k(J^l(T)/B^l)/C^k.$$

Let V' be another Γ-structure on V sufficiently close to V. Our first concern is to associate a section $\omega(V')$ over V of $\Theta(E_1)$ with V' in such a way that sections thus obtained are characterized by a certain equation. Unfortunately the writer does not know any canonical method for doing this. One way is to assume that V' can be joined to V by a family and to use exponential mappings, but the result seems to depend on how you join V' to V. Here we use a method depending on a choice of certain decompositions in jet spaces. To do this, we remark that the space of invertible l-jets of V, say $G^l(V)$, is an open subset of $J^l(J^0(V), V, \alpha)$ and the set of l-jets (for $l \geqq l_1$) belonging to $\Gamma(V)$ is a submanifold, say $A^l = J^l(\Gamma(V))$, of $G^l(V)$, i. e. $\Gamma(V)$ is the set of integrals of the differential equation A^l. We can form the coset manifold $G^l(V)/A^l$ by means of the composition of jets, elements of which are cosets $(A^l \cap \alpha^{-1}(p)) \circ X$ where $X \in G^l(V)$ with $p = \beta(X)$. $(G^l(V)/A^l, V, \alpha)$ is a fibered manifold. Denote by $I^l(p)$ the l-jet of identity mapping at p of V and by $'I^l(p)$ the image of $I^l(p)$ by the canonical projection of $G^l(V)$ to $G^l(V)/A^l$. Set

$$I^l(V)=\cup\,\{I^l(p)\,;p\in V\}\text{ and }'I^l(V)=\cup\,\{'I^l(p)\,;p\in V\}.$$

Similarly we define $O^l(p)$, $'O^l(p)$, $O^l(V)$, and $'O^l(V)$ using l-jets of the zero vector field and the factor vector bundle $E_1=J^l(T)/B^l$. Since local vector fields can be considered as the differentials of one-parameter family of local transformations at the identity mapping, there is a canonical isomorphism τ'' of the tangent vector space to $G^l(V)\cap\alpha^{-1}(p)$ at $I^l(p)$ with $J^l(T)$ over p. τ'' sends the tangent vector space to $A^l\cap\alpha^{-1}(p)$ at $I^l(p)$ to B^l over p. Therefore τ'' induces an isomorphism τ' of the tangent vector space to the fiber of $(G^l(V)/A^l, V,\alpha)$ at $'I^l(p)$ with E_1 over p. Therefore we can find an open neighborhood U of $'I^l(V)$ in $G^l(V)/A^l$, an open neighborhood $\mathfrak{U}$ of $'O^l(V)$ in E_1, and a homeomorphism τ of (U,V,α) with $(\mathfrak{U},V,\alpha)$ which induces the identity mapping on the base space V and such that $d\tau$ is an extension of the above τ'. τ induces a homeomorphism $j^k(\tau)$ of $J^k(U)=J^k(U,V,\alpha)$ onto $J^k(\mathfrak{U})=J^k(\mathfrak{U},V,\alpha)$. $J^k(U)$ and $J^k(\mathfrak{U})$ are open sets in $J^k(G^l(V)/A^l)$ and $J^k(E_1)$, respectively. We constructed the vector sub-bundle C^k of $J^k(E_1)$, which characterizes $j^k(D(\Theta(T)))$ ((1) and Theorem 2). By the same construction we find without much difficulty a locally closed submanifold H^k of $J^k(G^l(V)/A^l)$ such that ϕ is a cross-section of H^k if and only if ϕ is sufficiently close to $j^k('I^l)$ and $\phi=j^k(j^l(f)\bmod A^l)$, where $'I^l$ is the cross-section $p\to{}'I^l(p)$ and where f is a local transformation of V sufficiently close to the identity mapping. We can assume that H^k is in $J^k(U)$. Clearly $d(j^k(\tau))$ induces the homeomorphism of the tangent vector space to the fibre of (H^k,V,α) at $j_p{}^k('I^l)$ with the fiber in C^k over p, i.e. $j^k(\tau)(H^k\cap\alpha^{-1}(p))$ and $C^k\cap\alpha^{-1}(p)$ has the same tangent vector space at $O'(p)$. Taking a complementary vector sub-bundle K of C^k in $J^k(E_1)$ and denoting by π_c the canonical projection to C^k of $J^k(E_1)$ with the kernel K, we find that π_c induces an homeomorphism of $j^k(\tau)(H^k)$ onto an open neighborhood of the set of zeros in C^k and induces the identity mapping on each tangent vector space at the zeros, provided we choose sufficiently small H^k. Hence we can find a mapping r of C^k into K such that the following holds:

(i) r is fibers preserving and induces the identity mapping on the base space,

(ii) X in $J^k(E_1)$ sufficiently near to a zero-jet is in $j^k(\tau)(H^k)$ if and only if $X-r(\pi_c(X))$ is in C^k,

(iii) $r(t\xi)=0\pmod{t^2}$, where t is a real parameter and ξ is any element in C^k.

Now we fix such τ, K, and r once for all. Going back to a T-structure V', we take a sufficiently fine covering U_i of V. Choose coordinates x_i on U_i of the Γ-structure V and y_i on U_i of the Γ-structure V'. We assume that V' is sufficiently closed to V so that the following arguments go through: Take a covering V_i of V such that the closure of V_i is in U_i. We may assume that $y_i = x_i \circ g_i$ on V_i where g_i is a homeomorphism of V_i into U_i. .In the subsequent discussions we have to shrink V_i a little more. But we will omit mentioning it. Then, since V' is a Γ-structure, it follows that

$$g_i \circ g_j^{-1} \in \Gamma(V). \tag{2}$$

Now set

$$\omega_i = \tau(\rho_A(j^1(g_i)))$$

where ρ_A is the canonical projection of $G^1(V)$ to $G^1(V)/A^1$. ω_i is a cross-section of E_1 over V_i. Since (2) implies that $j^1(g_i) = a_{ij} \circ j^1(g_j)$ with $a_{ij} \in A^1$, it follows that $\omega_i = \omega_j$ on $V_i \cap V_j$. Thus ω_i determines a cross-section $\omega(V')$ of E_1 over V. By the construction $\omega(V')$ is independent of the choice of coordinates and the covering.

We will next characterize sections of E_1 obtained by the above method from Γ-structures. Let ω be a section of E_1 over V sufficiently close to the zero-section. Then there is a Γ-structure V' on V such that $\omega = \omega(V')$ if and only if we can find a system $\{g_i\}$ of homeomorphism g_i of V_i into U_i such that $\tau^{-1}(\omega)$ on V_i is equal to $\rho_A(j^1(g_i))$, because the fact that $\rho_A(j^1(g_i)) = \rho_A(j^1(g_j))$ on $V_i \cap V_j$ implies $g_i \circ g_j^{-1} \in \Gamma(V)$. The condition is equivalent to say that ω is an integral of the differential equation $j^k(\tau)(H^k)$. Then by the property (ii) of r, the condition is equivalent to the condition:

$$j^k(\omega) - r(\pi_c(j^k(\omega))) \in C^k. \tag{3}$$

Now $\rho_c \circ r \circ \pi_c \circ j^k$, where ρ_c is the canonical projection of $J^k(E_1)$ to $E_2 = J^k(E_1)/C^k$, defines a differential mapping (of order k) of $\mathfrak{S}(E_1)$ to $\mathfrak{S}(E_2)$ which we denote by R:

$$R = \rho_c \circ r \circ \pi_c \circ j^k.$$

Since $D = \rho_c \circ j^k$, the condition (3) is equivalent to the condition $D(\omega) - R(\omega) = 0$. Thus together with the property (iii) of r we have the following:

PROPOSITION 2.1. *Let ω be a cross-section of E_1 over V such that $j^k(\omega)$ is sufficiently close to the zero section. Then there is a Γ-structure V' on V such that $\omega = \omega(V')$ if and only if*

$$D(\omega) - R(\omega) = 0 \tag{4}$$

where D and R are differential operator of order k, D is linear and $R(t\omega) = 0$ $(\mathrm{mod}\, t^2)$.

Thus the problem of local deformations of Γ-structure V is reduced to the problem of solving the differential equation (4). At present we do not know any effective way of solving the equation for general Γ. However the much of the known argument in the case of complex analytic structure can be carried over to Γ-structures provided V is elliptic in the following sense: For a suitable choice of l and k such that the above construction can be carried out and Proposition 2.1 holds, we can find a Riemannian metric in V and positive definite metrics along the fibers of the vector bundles T, $J^l(T)/B^l$, and $J^k(J^l(T)/B^l)/C^k$ in such a way that the differential mapping $DD^* + D^*D$ of $\Theta(E_1)$ into itself is elliptic, where D^* is the adjoint operator of D with respect to the above metrics. Following the known argument we have

THEOREM 4. *Let V be an elliptic Γ-structure on a compact manifold V. Denote by $\Theta(V)$ the sheaf of germs of local vector fields belonging to V. Then the space of infinitesimal deformations of V, which is equal to $H^1(V, \Theta(V))$, is of finite dimension. If $H^1(V, \Theta(V)) = 0$, any family of deformations of V is locally trivial at the reference point. If $H^2(V, \Theta(V)) = 0$, there is a family of deformations of V such that its space of infinitesimal deformations at the reference point is equal to $H^1(V, \Theta(V))$.*

COLUMBIA UNIVERSITY.

REFERENCES.

[1] E. Cartan, *Les systèmes différentiels extérieurs et leur applications géometriques*, Paris, 1945.

[2] K. Kodaira, "On deformations of some complex pseudogroup structures," *Annals of Mathematics*, vol. 71 (1960), pp. 224-302.

[3] K. Kodaira and D. C. Spencer, "Multifoliate structures," *Annals of Mathematics*, vol. 71 (1960), pp. 43-76.

[4] M. Kuranishi, *Lectures on exterior differential systems*, Tata Institute of Fundamental Research, Bombay, 1962.

[5] D. C. Spencer, "Deformation of structures on manifolds defined by transitive continuous pseudogroups," *Annals of Mathematics*, vol. 76 (1962), pp. 306-445.

ON SIEGEL MODULAR FORMS OF GENUS TWO (II).[1]

By JUN-ICHI IGUSA.

Introduction. The main subject we shall discuss in this second paper is the Siegel modular forms of genus two with levels. The method we used in the first paper [5] did not give sufficient information even for level two. Therefore the problem (raised by Grothendieck) whether modular varieties become non-singular or not for higher levels was beyond our reach. With some other applications in mind, we therefore investigated "theta-constants" as modular forms and proved, among others, a fundamental lemma in our recent paper [6]. Using the results in that paper, we shall show that *modular varieties of high levels do not have non-singular coverings even locally around their singular points.* Also we shall determine how $\Gamma_2(1)/\Gamma_2(2) = Sp(2, \mathbf{Z}/2\mathbf{Z})$ acts on the ring of modular forms $A(\Gamma_2(2))$ and obtain the characters of its action on the homogeneous parts $A(\Gamma_2(2))_k$ for $k = 0, 1, 2, \cdots$. In this way, we shall determine $A(\Gamma_2(1))$ reproducing our earlier structure theorem on $A(\Gamma_2(1))^{(2)}$. Furthermore, the *polynomial expressions* of the four basic Eisenstein series of level one by theta-constants and a known identity of this kind (between a certain Eisenstein series of level two and the eighth power of Riemann's theta-constant [1]) will be obtained. We note that this identity was previously obtained using the Siegel main theorem on quadratic forms.

1. Normality of $C[\theta_{\mathfrak{m}}\ \theta_{\mathfrak{n}}]$.

We shall use the same notation as in TC except that we have $g = 2$ in this paper (and may omit $g = 2$ for the sake of simplicity). Also we shall sometimes treat characteristics as line vectors instead of column vectors (as in the case $g = 1$). We recall that theta-constants $\theta_{\mathfrak{m}}(\tau)$ are defined by putting $z = 0$ in

$$\theta_{\mathfrak{m}}(\tau, z) = \sum_{p \in \mathbf{Z}^2} e[\tfrac{1}{2} \cdot {}^t(p + m'/2)\tau(p + m'/2) + {}^t(p + m'/2)(z + m''/2)].$$

Also we shall denote the function $\tau \to \theta_{\mathfrak{m}}(\tau)$ defined in the Siegel upper-half plane $\mathfrak{S}$ by $\theta_{\mathfrak{m}}$. There are ten even characteristics mod 2 and they are (0000), (0001), (0010), (0011), (0100), (0110), (1000), (1001), (1100), (1111).

Received October 11, 1963.

[1] This work was partially supported by the National Science Foundation.

The purpose of this section is to prove the normality of the graded ring $\boldsymbol{C}[\theta_{\mathrm{m}}]$ generated over $\boldsymbol{C}$ by the ten "theta-constants."

LEMMA 1. *Let K be a field* (of characteristic zero) *and let $K[X, T]$ be a ring of polynomials in ten letters X_{ij} for $1 \leqq i, j \leqq 3$ and T with coefficients in K. Then the following twenty-one quadratic polynomials*

$$\sum_{1 \leqq j \leqq 3} X_{i_1 j} X_{i_2 j} - \delta_{i_1 i_2} T^2, \qquad \sum_{1 \leqq i \leqq 3} X_{i j_1} X_{i j_2} - \delta_{j_1 j_2} T^2$$
$$(X_{i_1 j_1} X_{i_2 j_2} - X_{i_1 j_2} X_{i_2 j_1}) - X_{i_3 j_3} T,$$

in which the permutation $(i_1 i_2 i_3) \rightarrow (j_1 j_2 j_3)$ is even, generate a prime ideal in $K[X, T]$. Moreover, if (x, t) is a generic zero over K, the integral domain $K[x, t]$ is normal.

Proof. We observe that we have an identity of the form

$$\begin{aligned}\det(X_{ij}) - T^3 = {} & X_{11}((X_{22}X_{33} - X_{23}X_{32}) - X_{11}T) \\ & + X_{21}((X_{32}X_{13} - X_{33}X_{12}) - X_{21}T) \\ & + X_{31}((X_{12}X_{23} - X_{13}X_{22}) - X_{31}T) \\ & + ((X_{11})^2 + (X_{21})^2 + (X_{31})^2 - T^2)T.\end{aligned}$$

Hence the first part becomes a special case of a known theorem and a proof can be found in Weyl [11, pp. 144-7]. We shall prove the second part. The main point is that every element of $K[x, t]$ can be written uniquely as a linear combination of monomials

$$x_{i_1 j_1} \cdots x_{i_p j_p} t^e \quad i_1 \leqq \cdots \leqq i_p, \; j_1 \geqq \cdots \geqq j_p$$

with coefficients in K in which x_{i1} and x_{3j} appear at most once. In order to make the argument clear, we consider the set of pairs of non-negative integers and introduce a lexicographic order in it. This set is well ordered. Then to each monomial $X_{i_1 j_1} \cdots X_{i_p j_p} T^e$ of $K[X, T]$, we associate the following element $(p, \sum i_\alpha + (3 - j_\alpha))$ of the well-ordered set as its "weight." Now we introduce three operations in $K[X, T]$. The first operation is

$$X_{i_1 j_1} X_{i_2 j_2} \rightarrow X_{i_1 j_2} X_{i_2 j_1} + X_{i_3 j_3} T$$

whenever both $i_1 < i_2$ and $j_1 < j_2$. The second and the third operations are

$$X_{i_1 1} X_{i_2 1} \rightarrow \delta_{i_1 i_2} T^2 - \sum_{j > 1} X_{i_1 j} X_{i_2 j}$$
$$X_{3 j_1} X_{3 j_2} \rightarrow \delta_{j_1 j_2} T^2 - \sum_{i < 3} X_{i j_1} X_{i j_2}.$$

We observe that, by the first operation each monomial is replaced by a sum of two monomials one of the same weight and another of a smaller weight. On the other hand, by the two other operations each monomial is replaced by a sum of two or three monomials of smaller weights. If a polynomial in X_{ij}, T is given, we apply the first operation repeatedly (as long as it can be applied) and we get a new polynomial which is a linear combination of monomials $X_{i_1j_1}\cdots X_{i_pj_p}T^e$ in which $i_1 \leqq \cdots \leqq i_p$, $j_1 \geqq \cdots \geqq j_p$. The new polynomial may be called "properly ordered." We then apply two other operations and order properly the new polynomial so obtained. Because weights decrease by the second and the third operations, the process necessarily stops after a finite number of times. In this way, we get a properly ordered polynomial each monomial of which contains X_{i1} and X_{3j} at most once. If we replace X_{ij} by x_{ij} and T by t, the three operations reduce to modifying expressions of the element of $K[x,t]$. Thus every element of $K[x,t]$ can be written in the way stated before. We shall prove its uniqueness. Suppose that we have a relation of the form

$$P_0(x_{ij}) + P_1(x_{ij})t + \cdots + P_n(x_{ij})t^n = 0$$

in which $P_e(x_{ij})t^e$ is a linear combination of the said monomials for $0 \leqq e \leqq n$. Also the relation is homogeneous in x_{ij} and t. We can assume that $P_0(x_{ij})$ is different from zero. Let (x_1, x_2, x_3) and (y_1, y_2, y_3) be two independent generic points over K of a quadratic cone defined by $X_1^2 + X_2^2 + X_3^2 = 0$. Then $(x_{ij}, t) \rightarrow (x_iy_j, 0)$ is a specialization over K, hence $P_0(x_iy_j) = 0$. We shall show that this implies $P_0(x_{ij}) = 0$. We know that $P_0(X_{ij})$ is a linear combination of $X_{i_1j_1}\cdots X_{i_nj_n}$ with coefficients c_{ij} say. If we introduce three more letters Y_1, Y_2, Y_3, the condition $P_0(x_iy_j) = 0$ means that

$$\sum c_{ij}X_{i_1}\cdots X_{i_n}Y_{j_1}\cdots Y_{j_n}$$

is contained in the ideal of $K[X_1, X_2, X_3, Y_1, Y_2, Y_3]$ generated by

$$X_1^2 + X_2^2 + X_3^2 \text{ and } Y_1^2 + Y_2^2 + Y_3^2.$$

On the other hand, neither X_3 nor Y_1 appears twice in the products $X_{i_1}\cdots X_{i_n}Y_{j_1}\cdots Y_{j_n}$. Since such monomials are linearly independent over K modulo the ideal in question, we get $c_{ij} = 0$ and $P_0(x_{ij}) = 0$. But this is a contradiction. This completes the proof of the key point.

Now, for the proof of the normality of $K[x,t]$, we can assume that K is algebraically closed or at least contains the primitive fourth root of unity i. Then there exists an automorphism of $K[x,t]$ over K permuting any x_{ij} with t (multiplied by a fourth root of unity). In fact, since the permutation of any two lines or columns of the 3×3 matrix (x_{ij}) followed by the change of signs of any line or column gives rise to an automorphism of $K[x,t]$ over K, we have only to show that x_{13}, for instance, and $-t$ are interchangeable. This we see by verifying that $t \to -x_{13}$ and

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} \to \begin{pmatrix} ix_{11} & ix_{12} & -t \\ -x_{33} & x_{31} & ix_{23} \\ x_{22} & -x_{21} & ix_{33} \end{pmatrix}$$

give rise to an automorphism of $K[x,t]$ over K. (The meaning of this automorphism will become clear in the next section.) This being remarked, we consider (x,t) as a homogeneous generic point over K of a variety embedded in a projective space. Then its affine open set $t \neq 0$ is the variety of the special orthogonal group, and it is non-singular. Therefore, by the above remark, the projective variety is non-singular. The rest of the proof is similar to our proof of the artihmetic normality of the Grassmann variety [4]. We observe that the conductor of $K[x,t]$ is irrelevant in the sense it contains a power of the maximal ideal of $K[x,t]$ generated by x_{ij} and t. Consequently, the graded ring $K[x,t]$ and its normalization coincide except for homogeneous parts of lower degrees. Suppose that they coincide at degree n but not at degree $n-1$. Let ξ be an element of degree $n-1$ in the normalization which is not contained in $K[x,t]$. Then $t\xi$ is in $K[x,t]_n$, hence we have

$$t\xi = P_0(x_{ij}) + P_1(x_{ij})t + \cdots + P_n(x_{ij})t^n$$

in which $P_e(x_{ij})t^e$ is a linear combination of the previously explained monomials (of degree n) for $0 \leqq e \leqq n$. Since ξ is not contained in $K[x,t]$, we have $P_0(x_{ij}) \neq 0$. Apply the specialization $(x_{ij}, t) \to (x_i y_j, 0)$ over K as before. Then, since ξ is integral over $K[x,t]$, we have $t\xi \to 0$, hence $P_0(x_i y_j) = 0$ and $P_0(x_{ij}) = 0$. But this is a contradiction. q. e. d.

LEMMA 2. *Let R be a* (noetherian) *normal domain in which 2 is a unit and let $a_1, a_2, \cdots, a_n$ be non-units of R such that $Ra_i + Ra_j$ has rank two for $i \neq j$. Suppose further that the quotient ring R/Ra_i has no nilpotent element. Let b_i be a square root of a_i for $i = 1, 2, \cdots, n$. Then $R[b_1, b_2, \cdots, b_n]$ is a normal domain.*

Proof. Put $R_1 = R[b_1]$. If we can show that R_1 and $a_2, \cdots, a_n$ satisfy similar conditions as R and $a_1, a_2, \cdots, a_n$, we can apply an induction on n. Let K be the field of fractions of R. Then $K(b_1)$ is clearly a quadratic extension of K and R_1 is the normalization of R in $K(b_1)$. Moreover 2 is a unit and $a_2, \cdots, a_n$ are non-units of R_1. The "going-down" theorem of Cohen-Seidenberg [cf. 8, pp. 31-32] asserts that $R_1a_i + R_1a_j$ has rank two for $i \neq j$. We shall show that R_1/R_1a_i has no nilpotent element for $i = 2, \cdots, n$. Otherwise there exists an element c of R_1 which itself is not but its square is contained in R_1a_2 say. Write c in the form $p + qb_1$ with p, q in R. Then we have $c^2 = (p^2 + q^2a_1) + 2pqb_1 = (p' + q'b_1)a_2$ with some p', q' in R. This implies $p^2 + q^2a_1 = p'a_2$, $2pq = q'a_2$, hence $p^2 + q^2a_1$ and $2pq$ are in Ra_2. Since Ra_2 is an intersection of prime ideals not containing a_1, this implies that p and q are in Ra_2, hence c is in R_1a_2. But this is a contradiction. The lemma is thus proved.

LEMMA 3. *Let K, x_{ij}, t be as in Lemma* 1 *and let $(x_{ij})^{\frac{1}{2}}$, $t^{\frac{1}{2}}$ be square roots of x_{ij}, t for $1 \leqq i, j \leqq 3$. Then $K[x^{\frac{1}{2}}, t^{\frac{1}{2}}]$ is a normal domain.*

Proof. We shall show that the ten principal ideals $K[x,t]x_{ij}$, $K[x,t]t$ are prime and distinct. First of all, a part of the proof of Lemma 1 shows that $K[x,t]t$ is the kernel of the homomorphism of $K[x,t]$ associated with the specialization $(x_{ij}, t) \to (x_iy_j, 0)$ over K. Therefore $K[x,t]t$ is a prime ideal. Since we know that any x_{ij} is interchangeable with t (multiplied by a fourth root of unity), we see that $K[x,t]x_{ij}$ is a prime ideal also. Since under the specialization $(x_{ij}, t) \to (x_iy_j, 0)$ over K, no x_{ij} is specialized to zero, the prime ideal $K[x,t]t$ is different from any other prime ideal $K[x,t]x_{ij}$. Therefore, no two of the ten prime ideals are same. The rest follows from Lemma 2 by taking $K[x,t]$ as R and the ten elements x_{ij}, t as $a_1, a_2, \cdots, a_n$.

Now the preparation is made to prove the following theorem:

THEOREM 1. *The graded rings $\boldsymbol{C}[(\theta_{\mathfrak{m}})^2]$, $\boldsymbol{C}[\theta_{\mathfrak{m}}]$ and $\boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}]$ are all normal, hence $A(\Gamma(4,8)) = \boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}]$.*

Proof. If we use Riemann's theta-formula, which we explained in TC, Section 4, we can verify that

$$t = -(\theta_{0110})^2$$

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} = \begin{pmatrix} (\theta_{1001})^2 & -i(\theta_{1000})^2 & (\theta_{0100})^2 \\ i(\theta_{0001})^2 & (\theta_{0000})^2 & i(\theta_{1100})^2 \\ (\theta_{0010})^2 & -i(\theta_{0011})^2 & (\theta_{1111})^2 \end{pmatrix}.$$

define a zero of the ideal introduced in Lemma 1. Since the field $\boldsymbol{C}(\theta_{\mathfrak{m}})$ has dimension four over $\boldsymbol{C}$, we actually have a generic zero over $\boldsymbol{C}$ and the notation is justified. Therefore $\boldsymbol{C}[(\theta_{\mathfrak{m}})^2] = \boldsymbol{C}[x,t]$ is normal by Lemma 1. Hence $\boldsymbol{C}[\theta_{\mathfrak{m}}] = \boldsymbol{C}[x^{\frac{1}{2}}, t^{\frac{1}{2}}]$ is normal by Lemma 3. Since we have $\boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}] = \boldsymbol{C}[\theta_{\mathfrak{m}}]^{(2)}$, therefore $\boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}]$ is also normal. Then, TC, Theorem 5 implies that $\boldsymbol{C}[\theta_{\mathfrak{m}}\theta_{\mathfrak{n}}]$

coincides with the ring of modular forms $A(\Gamma(4,8))$ belonging to the congruence group $\Gamma(4,8)$. This completes the proof.

2. Going-down process. Since we know the ring $A(\Gamma(4,8))$, by taking invariant subrings successively, we can determine $A(\Gamma(4))$, $A(\Gamma(2,4))$, $A(\Gamma(2))$ and finally $A(\Gamma(1))$. This procedure is explained in TC, Section 6. The ring $A(\Gamma(4))$ is generated over $\boldsymbol{C}$ by the ten $(\theta_{\mathfrak{m}})^2$, fifteen products $\theta_{\mathfrak{m}_1}\theta_{\mathfrak{m}_2}\,\theta_{\mathfrak{m}_3}\theta_{\mathfrak{m}_4}$ of distinct $\theta_{\mathfrak{m}}$'s satisfying

$$\sum_{1\leqq\alpha\leqq 4} m_{\alpha i}' \equiv \sum_{1\leqq\alpha\leqq 4} m_{\alpha i}'' \equiv 0 \bmod 2,$$

fifteen "complementary" products $\theta_{\mathfrak{m}_1}\theta_{\mathfrak{m}_2}\,\theta_{\mathfrak{m}_3}\theta_{\mathfrak{m}_4}\,\theta_{\mathfrak{m}_5}\theta_{\mathfrak{m}_6}$ and by

$$\theta = \prod_{\mathfrak{m}\text{ even, mod } 2} \theta_{\mathfrak{m}}.$$

The ring $A(\Gamma(2,4))$ is generated over $\boldsymbol{C}$ by $(\theta_{\mathfrak{m}}\theta_{\mathfrak{n}})^2$ and θ, i.e. $A(\Gamma(2,4)) = \boldsymbol{C}[x,t]^{(2)}[\theta]$. In particular, *the modular variety* proj. $A(\Gamma(2,4))$ *is non-singular.* The ring $A(\Gamma(2))$ is generated over $\boldsymbol{C}$ by the ten $(\theta_{\mathfrak{m}})^4$, the fifteen products $(\theta_{\mathfrak{m}_1}\theta_{\mathfrak{m}_2}\,\theta_{\mathfrak{m}_3}\theta_{\mathfrak{m}_4})^2$, the fifteen complementary products $(\theta_{\mathfrak{m}_1}\theta_{\mathfrak{m}_2}\,\theta_{\mathfrak{m}_3}\theta_{\mathfrak{m}_4}\,\theta_{\mathfrak{m}_5}\theta_{\mathfrak{m}_6})^2$ and by θ. If we cross out from these products (except θ) those which are not of the form $x_{i_1j_1}\cdots x_{i_pj_p}t^e$ where $i_1\leqq\cdots\leqq i_p$, $j_1\geqq\cdots\geqq j_p$ and in which x_{i1} and x_{3j} appear at most once, we get t^2, $(x_{12})^2$, $(x_{13})^2$, $(x_{22})^2$, $(x_{23})^2$, $x_{13}x_{22}x_{31}t$ and θ. Since we have

$$\begin{aligned}2x_{13}x_{22}x_{31}t = (t^2 - (x_{12})^2 - (x_{13})^2 - (x_{22})^2 - (x_{23})^2)t^2 \\ + (x_{12}x_{23})^2 - (x_{13}x_{22})^2,\end{aligned}$$

actually the element $x_{13}x_{22}x_{31}t$ is redundant. Thus $A(\Gamma(2))$ is generated over $\boldsymbol{C}$ by θ and by

$$\begin{aligned} &y_0 = t^2, \quad y_1 = (x_{13})^2, \quad y_2 = (x_{22})^2 \\ &y_3 = (x_{12})^2 - t^2, \quad y_4 = (x_{23})^2 - t^2. \end{aligned}$$

Furthermore these elements are related as follows

$$\begin{aligned} &(y_0y_1 + y_0y_2 + y_1y_2 - y_3y_4)^2 - 4y_0y_1y_2(y_0 + y_1 + y_2 + y_3 + y_4) = 0 \\ &\theta^2 = \tfrac{1}{4}\cdot(y_0y_1 + y_0y_2 + y_1y_2 - y_3y_4) \\ &\qquad \cdot(2y_0y_1y_2 + y_0y_1y_3 + y_0y_1y_4 + y_0y_2y_3 + y_0y_2y_4 \\ &\qquad + 2y_0y_3y_4 + y_1y_2y_3 + y_1y_2y_4 + 2y_1y_3y_4 + 2y_2y_3y_4 \\ &\qquad + y_3^2y_4 + y_3y_4^2). \end{aligned}$$

In particular *the modular variety* proj. $A(\Gamma(2))$ *is a quartic hypersurface* (defined by the above equation) *in the four dimensional projective space.*[2] The going-down process from $A(\Gamma(2))$ to $A(\Gamma(1))$ is more involved. We have to know how $\Gamma(1)/\Gamma(2)$ acts on $A(\Gamma(2))$ and, in particular, how the characters of its action on the homogeneous parts $A(\Gamma(2))_k$ decompose into simple characters for $k=0,1,2,\cdots$.

We first observe that $\Gamma(1)/\Gamma(2)=Sp(2,\mathbf{Z}/2\mathbf{Z})$ is isomorphic to the symmetric group π_6 of permutations on six letters. The isomorphism can be given as follows. Suppose that M is an element of $\Gamma(1)$ composed of a, b, c, d and $\mathfrak{m}$ an arbitrary characteristic. Then $M\cdot\mathfrak{m}$ was defined as

$$M\cdot\mathfrak{m}=\begin{pmatrix} d & -c \\ -b & a\end{pmatrix}\mathfrak{m}+\begin{pmatrix}(c^td)_0\\(a^tb)_0\end{pmatrix}.$$

If we consider characteristics mod 2, the group $\Gamma(1)/\Gamma(2)$ operates on the set of odd (as well as even) characteristics. There are six odd characteristics mod 2 and they are (0101), (0111), (1010), (1011), (1101), (1110). Since the only element of $Sp(2,\mathbf{Z}/2\mathbf{Z})$ keeping these characteristics fixed mod 2 is the identity, we have a monomorphism of $Sp(2,\mathbf{Z}/2\mathbf{Z})$ to π_6. Since the orders of the two groups are same, we have an isomorphism.

This being remarked, we shall determine the character of the representation of $Sp(2,\mathbf{Z}/2\mathbf{Z})$ on $A(\Gamma(2))_k$. We observe that we have

$$\begin{aligned}\dim_{C}A(\Gamma(2))_k&=\binom{k/2+4}{4}-\binom{k/2}{4}\\&=1/12\cdot(k^3+3k^2+14k+12)\end{aligned}$$

for $k=0,2,4,\cdots$ and

$$\dim_{C}A(\Gamma(2))_k=\dim_{C}A(\Gamma(2))_{k-5}$$

for $k=1,3,5,\cdots$. In particular the dimensions are 1, 0, 5, 0, 15, 1, $\cdots$ for $k=0$, 1, 2, 3, 4, 5, $\cdots$. We shall determine how $Sp(2,\mathbf{Z}/2\mathbf{Z})$ operates on $A(\Gamma(2))_2$. This can be done using TC, Theorem 6. We do it as follows. We call the six odd characteristics mod 2 simply $1,2,\cdots,6$. We know in general that the symmetric group π_g on g letters is generated by (12) and $(12\cdots g)$. The element M of $\Gamma(1)$ which gives rise to (12) satisfies

$$M\equiv\begin{pmatrix}1&0&1&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\bmod 2.$$

---

[2] If S is a graded integral domain (of finite type over C), we have proj. $S=$ proj. $S^{(d)}$ for every positive integer d. We are using this fact for $d=2$ in the above two cases.

Also the element M of $\Gamma(1)$ which gives rise to (123456) satisfies

$$M \equiv \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ -1 & 1 & 0 & 1 \end{pmatrix} \bmod 2.$$

We shall denote the elements of $\Gamma(1)$ standing on the right sides by M_{12} and $M_{1\cdots 6}$ respectively. We have seen in TC, Section 2 that the transformation law of theta-functions implies

$$\theta_{M\cdot\mathfrak{m}}(M\cdot\tau) = \varkappa(M)e(\phi_{\mathfrak{m}}(M))\det(c\tau+d)^{\frac{1}{2}}\theta_{\mathfrak{m}}(\tau)$$

for all M in $\Gamma(1)$ in which

$$\phi_{\mathfrak{m}}(M) = -\tfrac{1}{8}\cdot({}^tm'{}^tbdm' + {}^tm''{}^tacm'' - 2{}^tm'{}^tbcm'' \\ - 2^t(a^tb)_0(dm' - cm'')).$$

We can determine $\varkappa(M_{12})$ and $\varkappa(M_{1\cdots 6})$ as in TC, Section 3. In particular we have

$$\varkappa(M_{12})^2 = 1, \qquad \varkappa(M_{1\cdots 6})^2 = e(\tfrac{1}{4}).$$

Now, in order to obtain a concrete matricial representation of $Sp(2,\boldsymbol{Z}/2\boldsymbol{Z})$ acting on the five dimensional complex vector space $A(\Gamma(2))_2$, we arrange $y_0, y_1, \cdots, y_4$ in a column vector $\mathfrak{y}$. Applying the formula we just copied from TC, we then get

$$M^{-1}\cdot\mathfrak{y} = \rho(M)\mathfrak{y}$$

with

$$\rho(M_{12}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & -1 & 0 \\ -1 & -1 & 0 & 0 & -1 \end{pmatrix}, \quad \rho(M_{1\cdots 6}) = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix}.$$

We also have

$$M_{12}{}^{-1}\cdot\theta = M_{1\cdots 6}{}^{-1}\cdot\theta = -\theta.$$

We shall combine these results with the known theory of representations of symmetric groups [3, 11].

We know in general that conjugacy classes of π_g correspond to partitions of g. In fact every element of π_g has a cycle expression and it determines a partition of g. Two elements of π_g are conjugate if and only if they correspond to the same partition of g. On the other hand, simple characters of π_g also correspond to partitions of g. In fact each partition of g determines a primitive idempotent of the group ring of π_g over $\boldsymbol{Q}$ and it gives a simple character of π_g and conversely. Now the number 6 has the following eleven partitions

$$\begin{aligned}6&=5+1=4+2=4+1+1=3+3=3+2+1\\&=3+1+1+1=2+2+2=2+2+1+1\\&=2+1+1+1+1=1+1+1+1+1+1.\end{aligned}$$

We arrange eleven conjugacy classes and simple characters of π_6 in this order (one from top to bottom another from left to right) and write them as follows:

1	-1	0	1	0	0	-1	0	0	1	-1	120
1	0	-1	0	0	1	0	0	-1	0	1	144
1	-1	1	0	-1	0	0	-1	1	-1	1	90
1	1	-1	0	-1	0	0	1	1	-1	-1	90
1	-1	0	1	2	-2	1	2	0	-1	1	40
1	0	0	-1	1	0	1	-1	0	0	-1	120
1	2	0	1	-1	-2	1	-1	0	2	1	40
1	-1	3	-2	-3	0	2	3	-3	1	-1	15
1	1	1	-2	1	0	-2	1	1	1	1	45
1	3	3	2	1	0	-2	-1	-3	-3	-1	15
1	5	9	10	5	16	10	5	9	5	1	1.

The last column which is separated from the main part indicates the numbers of elements in the conjugacy classes.

The table shows that the representation ρ is irreducible and its character is the character χ_{222} corresponding to the partition $6=2+2+2$. Using the representation ρ, we operate $Sp(2,\boldsymbol{Z}/2\boldsymbol{Z})$ on the graded ring $\boldsymbol{C}[Y]$ of polynomials in five letters $Y_0, Y_1, \cdots, Y_4$ each of degree two with coefficients in $\boldsymbol{C}$. If we denote by Ψ_k the character of the representation on $\boldsymbol{C}[Y]_k$, for every M in $Sp(2,\boldsymbol{Z}/2\boldsymbol{Z})$, we have the following identity

$$\sum_{0\leqq k<\infty}\Psi_{2k}(M)t^k=1/\det(1_5-\rho(M)t)$$

of formal power-series in t. We then observe that $A(\Gamma(2))^{(2)}$ is the quotient ring of $\boldsymbol{C}[Y]$ by a principal ideal generated by

$$(Y_0Y_1+Y_0Y_2+Y_1Y_2-Y_3Y_4)^2-4Y_0Y_1Y_2(Y_0+Y_1+\cdots+Y_4).$$

This polynomial (in fact both

$$Y_0Y_1+Y_0Y_2+Y_1Y_2-Y_3Y_4 \text{ and } Y_0Y_1Y_2(Y_0+Y_1+\cdots+Y_4))$$

are invariant by the operations of $Sp(2,\mathbf{Z}/2\mathbf{Z})$. Therefore, if X_k is the character of the action of $Sp(2,\mathbf{Z}/2\mathbf{Z})$ on $A(\Gamma(2))_k$, we have

$$\sum_{0\leqq k<\infty} \mathrm{X}_{2k}t^k = (1-t^4)\sum_{0\leqq k<\infty} \Psi_{2k}t^k.$$

Moreover, if k is odd, we have $A(\Gamma(2))_k = A(\Gamma(2))_{k-5}\cdot\theta$ and the element θ has the property $M\cdot\theta = \epsilon(M)\theta$, in which $\epsilon(M) = \pm 1$ according as M corresponds to even or odd permutation of π_6. We note that ϵ is the simple character corresponding to the partition $6 = 1+1+1+1+1+1$. In this way, we get the following definitive result:

THEOREM 2. *The characters* X_k *of the representations of* $Sp(2,\mathbf{Z}/2\mathbf{Z})$ *on* $A(\Gamma(2))_k$ *are given by*

$$\sum_{0\leqq k<\infty} \mathrm{X}_k(M)t^k = (1+\epsilon(M)t^5)(1-t^8)/\det(1_5-\rho(M)t^2).$$

A complete table of $\det(1_5-\rho(M)t)$ can be obtained very easily and it is as follows:

$(1-t)^2(1+t)(1+t+t^2)$	(123456)
$1-t^5$	(12345)
$(1-t)(1+t)^2(1+t^2)$	$(1234)(56)$
$(1-t)^2(1+t)(1+t^2)$	(1234)
$(1-t)^3(1+t+t^2)$	$(123)(456)$
$(1+t)(1-t+t^2)(1+t+t^2)$	$(123)(45)$
$(1-t)(1+t+t^2)^2$	(123)
$(1-t)^4(1+t)$	$(12)(34)(56)$
$(1-t)^3(1+t)^2$	$(12)(34)$
$(1-t)^2(1+t)^3$	(12)
$(1-t)^5$	(1).

The symbols standing on the right are representatives of conjugacy classes and the ordering is the same as before. We note that this table can be obtained knowing only that we have $\mathrm{trace}(\rho) = \chi_{222}$ and knowing the values of χ_{222}.

3. Going-down process (continued). We recall that, if G is an arbitrary finite group and X a character of G, the multiplicity m_χ of a simple character χ in X can be determined as

$$m_\chi = 1/\mathrm{ord}(G)\cdot\sum_{s\in G}\mathrm{X}(s)\chi(s^{-1}).$$

Also, if G is a symmetric group, we have $\chi(s^{-1})=\chi(s)$ for every s. Using the results in the previous section, therefore, we can decompose the character X_k into simple characters. In particular $\dim_C A(\Gamma(1))_k$ can be calculated as the multiplicity of the principal character in X_k and we get

$$\begin{aligned}\sum_{0\leqq k<\infty}\dim_C A(\Gamma(1))_k t^k &\\ &=(1+t^{35})/(1-t^4)(1-t^6)(1-t^{10})(1-t^{12})\\ &=1+t^4+t^6+t^8+2t^{10}+3t^{12}+\cdots.\end{aligned}$$

The calculation (which we are not writing down here) shows also that, if we denote by $\Gamma(1)_e$ the subgroup of $\Gamma(1)$ of index two defined by $\epsilon(M)=1$, we get

$$\begin{aligned}\sum_{0\leqq k<\infty}\dim_C A(\Gamma(1)_e)_k t^k &\\ &=(1+t^{30})/(1-t^4)(1-t^5)(1-t^6)(1-t^{12}).\end{aligned}$$

At any rate, since we know

$$\begin{aligned}\sum_{0\leqq k<\infty}\dim_C A(\Gamma_1(1))_k t^k &=1/(1-t^4)(1-t^6)\\ &=1+t^4+t^6+t^8+t^{10}+2t^{12}+\cdots,\end{aligned}$$

the ring $A(\Gamma(1))$ is generated over C by two modular forms of weights four, six and three cusp forms of weights ten, twelve, thirty-five and all are unique up to constant factors. We also remarked previously [5, Appendix] that the definition and elementary properties of Eisenstein series plus this information imply that, if we denote by ψ_k the Eisenstein series of weights k for $k=4$, $6,\cdots$, the modular forms of weights four, six are constant multiples of ψ_4, ψ_6 and the cusp forms of weights ten and twelve are constant multiples of $\psi_4\psi_6-\psi_{10}$ and $3^27^2(\psi_4)^3+2\cdot 5^3(\psi_6)^2-691\psi_{12}$. We shall construct these modular forms explicitly using theta-constants.

We shall denote $(\theta_m)^4$ corresponding to ten even characteristics mod 2 in the order of Section 1 by $(1),(2),\cdots,(10)$. Then a general observation in TC, Appendix shows that

$$\tfrac{1}{4}\cdot\sum_{1\leqq i\leqq 10}(i)^2=\tfrac{1}{8}\cdot\sum_{i<j}e(i,j)(i)(j)$$

is an element of $A(\Gamma(1))_4$ which is mapped to 1 by the square of the Siegel operator. Hence this is ψ_4. In order to obtain the element of degree six, we apply a similar method, i.e. the symmetrization of products of theta-constants. We observe that the symmetrization of $(1)^3$ and $(1)^2(2)$ are both zero. Therefore we shall consider the symmetrization of $(1)(2)(3)$. We observe that $\Gamma(1)$ does not operate transitively on the set of all triples of even characteristics (mod 2). In fact

$$e(i,j,k)=e(i,j)e(j,k)e(k,i)$$

is invariant under the operations of $\Gamma(1)$, and it is 1 for $(1,2,3)$ and is -1 for $(1,2,5)$. They are known as "syzygous" and "azygous" triples. At any rate, we can verify that $\Gamma(1)_e$, hence a fortiori $\Gamma(1)$, operates transitively on the sets of syzygous and azygous triples of even characteristics. We therefore apply symmetrizations with respect to $\Gamma(1)_e$ to $((1)(2)(3))^m$ and to $((1)(2)(5))^m$, in general, and put

$$(\mathrm{syzy})_m=\tfrac{1}{4}\cdot\sum_{\text{syzygous}}(\pm(i)(j)(k))^m$$

$$(\mathrm{azy})_m=\tfrac{1}{8}\cdot\sum_{\text{azygous}}(\pm(i)(j)(k))^m$$

for $m=1,2,3,\cdots$. The summation in $(\mathrm{syzy})_1$, for instance, is extended over sixty syzygous triples of which thirty have plus-sign and the remaining thirty have minus-sign. The summation in $(\mathrm{azy})_1$ is similar. We observe that $(\mathrm{syzy})_m$ for every m and $(\mathrm{azy})_m$ for even m are elements of $A(\Gamma(1))_{6m}$ while $(\mathrm{azy})_m$ for odd m belongs to $A(\Gamma(1)_e)_{6m}$ and not to $A(\Gamma(1))$ unless it is zero. Furthermore the square of the Siegel operator maps $(\mathrm{syzy})_m$ to 1, hence $(\mathrm{syzy})_1$ is just a different expression for ψ_6. As for the cusp form of weight ten, we have one such form, i. e. θ^2. Moreover the Fourier expansion of θ itself starts as follows

$$\theta\begin{pmatrix} w_1 & \epsilon \\ \epsilon & w_2 \end{pmatrix}=-2^7i\cdot e(w_1/2)e(w_2/2)(\pi\epsilon)+\cdots.$$

There is no ready-made cusp form of weight twelve. We therefore symmetrize $(1)^6$ and try to determine a linear combination of the symmetrized element and $(\psi_4)^3$, $(\psi_6)^2$ to get a cusp form. In the calculation, it is understood that we use the expressions of ψ_4, ψ_6 by theta-constants. This simple consideration works and we get a cusp form together with the following Fourier expansion

$$\Bigl(3^2\sum_{1\leqq i\leqq 10}(i)^6-2^2 11(\psi_4)^3+2^3(\psi_6)^2\Bigr)\begin{pmatrix} w_1 & \epsilon \\ \epsilon & w_2 \end{pmatrix}$$
$$=2^{15}3^4 11e(w_1)e(w_2)+\cdots.$$

Finally we shall obtain the cusp form of weight thirty-five. The formula for $\dim_C A(\Gamma(1)_e)_k$ shows that it is a product of θ and an element of $A(\Gamma(1)_e)_{30}$ not in $A(\Gamma(1))$. We just have one such element, i. e. $(\mathrm{azy})_5$ provided it is different from zero. We can verify easily that $(\mathrm{azy})_m$ are different from zero except for $m = 1, 3$. Furthermore $(\mathrm{azy})_5$ has the following Fourier expansion

$$(\mathrm{azy})_5\begin{pmatrix} w_1 & \epsilon \\ \epsilon & w_2 \end{pmatrix} = -2^{29}5^3 e(3w_1/2)e(3w_2/2)(e(w_1) - e(w_2)) + \cdots.$$

We shall summarize our results in the following way:

THEOREM 3. *The graded ring $A(\Gamma(1))$ is generated by ψ_4, ψ_6 and three cusp forms χ_{10}, χ_{12}, χ_{35} of weights ten, twelve, thirty-five with Fourier expansions of the form*

$$\chi_{10}\begin{pmatrix} w_1 & \epsilon \\ \epsilon & w_2 \end{pmatrix} = e(w_1)e(w_2)(\pi\epsilon)^2 + \cdots$$

$$\chi_{12}\begin{pmatrix} w_1 & \epsilon \\ \epsilon & w_2 \end{pmatrix} = e(w_1)e(w_2) + \cdots$$

$$\chi_{35}\begin{pmatrix} w_1 & \epsilon \\ \epsilon & w_2 \end{pmatrix} = e(2w_1)e(2w_2)(e(w_1) - e(w_2))(\pi\epsilon) + \cdots.$$

They are all expressed by theta-constants. Moreover the ideal of cusp forms is generated in $A(\Gamma(1))$ by χ_{10}, χ_{12}, χ_{35}.

Only the last statement needs a proof and it is as follows. Let J be the ideal of $A(\Gamma(1))$ generated by χ_{10}, χ_{12}, χ_{35}. Since a cusp form of odd weight is in $A(\Gamma(1))\chi_{35}$, it is in J. Therefore we have only to show that cusp forms of even weights can be written in the form $\psi\chi_{10} - \psi'\chi_{12}$ with ψ, ψ' in $A(\Gamma(1))^{(2)}$. Since $A(\Gamma(1))^{(2)}$ is just a ring of polynomials in the four independent variables ψ_4, ψ_6, χ_{10}, χ_{12} over $\boldsymbol{C}$, we have $\psi\chi_{10} - \psi'\chi_{12} = 0$ if and only if $\psi = \psi''\chi_{12}$, $\psi' = \psi''\chi_{10}$ with ψ'' in $A(\Gamma(1))^{(2)}$. Therefore, for every even positive integer k, the dimension of $(A(\Gamma(1))/J)_k$ is given by

$$\begin{aligned} &\mathrm{coeff}_{t^k}((1 - t^{10} - t^{12} + t^{22})/(1 - t^4)(1 - t^6)(1 - t^{10})(1 - t^{12})) \\ &\qquad = \mathrm{coeff}_{t^k}(1/(1 - t^4)(1 - t^6)), \end{aligned}$$

and this is the dimension of $(A(\Gamma_1(1)))_k$. Hence J_k is precisely the kernel of the Siegel operator $A(\Gamma(1))_k \to A(\Gamma_1(1))_k$, and the theorem is proved.

We note that $(\chi_{35})^2$ is in $A(\Gamma(1))^{(2)}$, hence it is a polynomial in ψ_4, ψ_6, χ_{10}, χ_{12} with coefficients in $\boldsymbol{C}$. This modular form is connected with the "skew-invariant" of binary sextics [cf. 5]. Also we make the following statement:

COROLLARY. *The graded ring $A(\Gamma(1)_e)$ is generated by ψ_4, ψ_6 and the three cusp forms θ, χ_{12}, $(\mathrm{azy})_5$ of weights five, twelve, thirty.*

Finally, we shall consider the ring of modular forms $A(\Gamma(1,2))$. Since

$\Gamma(1,2)$ is the stabilizer of $(1)=(\theta_{0000})^4$, it is a subgroup of $\Gamma(1)$ of index ten. Moreover the image of $\Gamma(1,2)$ in π_6 can be determined. It is the subgroup of π_6 generated by permutations of 1, 4, 6 (among themselves), permutations of 2, 3, 5 and by $(12)(34)(56)$. In particular, the numbers of elements in the image corresponding to various partitions of 6 in the previous section are in that order 12, 0, 18, 0, 4, 12, 4, 6, 9, 6, 1. Using this fact and the results of the previous section, we get

$$\begin{aligned}\sum_{0\leqq k<\infty} \dim_C A(\Gamma(1,2))_k t^k &= P(t)/(1-t^2)(1-t^4)(1-t^6)(1-t^{12})\\ &= 1+2t^4+2t^6+4t^8+5t^{10}+9t^{12}+\cdots,\end{aligned}$$

in which $P(t)=(1-t^2+t^4+t^8)+t^{19}(1+t^4-t^6+t^8)$. In particular the dimension of $A(\Gamma(1,2))_4$ is two. We know two linearly independent elements of $A(\Gamma(1,2))_4$ already, and they are $(1)^2=(\theta_{0000})^8$ and ψ_4. On the other hand, if we denote the Eisenstein series of weight four belonging to $\Gamma(1,2)$ simply by ψ_4', this is also an element of $A(\Gamma(1,2))_4$. Hence we get a relation of the form

$$\psi_4'=p(\theta_{0000})^8+q\psi_4$$

with some p, q in $\boldsymbol{C}$. If we symmetrize this relation with respect to $\Gamma(1)$, we get $4p+10q=4$. On the other hand, if we apply the square of the Siegel operator to the above relation, we get $p+q=1$. The two equations have a unique solution $(p,q)=(1,0)$. Going back to the definitions of ψ_4' and θ_{0000}, we therefore get the *Braun identity*

$$\sum_{c^t d\ \mathrm{even}} \det(c\tau+d)^{-4}=\Big(\sum_{p\in Z^2} e(\tfrac{1}{2}\cdot {}^t p\tau p)\Big)^8.$$

We can give a more systematic treatment using Fourier expansions of Eisenstein series and also we can discuss relations of Eisenstein series of higher weights.

4. **Singularities of modular varieties.** We recall that, if Γ is a subgroup of $Sp(2,\boldsymbol{R})$ containing $\Gamma(2,4)$ as a subgroup of finite index, the corresponding modular variety proj.$A(\Gamma)$ has proj.$A(\Gamma(2,4))$ as a (global) non-singular covering. Therefore, the singularities of proj.$A(\Gamma)$, in particular those of proj.$A(\Gamma(1))$ and of proj.$A(\Gamma(2))$, are not too bad. We shall show, however, that the singularities of proj.$A(\Gamma(n))$ are quite bad for $n\geqq 3$. Let ∞ be the 0-dimensional boundary component of $\mathfrak{S}$ defined by

$$\infty=\lim_{\eta\to\infty} i\eta 1_2.$$

The values of $\theta_{\mathfrak{m}}$ at ∞, which are the images of $\theta_{\mathfrak{m}}$ by the square of the Siegel operator, are simply 1 for $m' = (00)$ and zero otherwise. At any rate, the point ∞ has a unique image point in every modular variety.

THEOREM 4. *There is no regular local ring which is integral over the analytic local ring of the modular variety* $\operatorname{proj}.A(\Gamma(4))$ *at the image point of* ∞.

Before we start proving this theorem, we recall the footnote 2 saying in particular that, in the consideration of modular varieties we can restrict modular forms to those of even weights. Therefore we shall talk about rings of modular forms without adding the superscript (2). We need some lemmas.

LEMMA 4. *If we put*

$$u_1 = (\theta_{0100}/\theta_{0000})^2, \qquad u_2 = (\theta_{1000}/\theta_{0000})^3$$
$$u_3 = (\theta_{1100}/\theta_{0000})^2,$$

these three functions form a set of local parameters of $\operatorname{proj}.A(\Gamma(2,4))$ *at the image point of* ∞.

Proof. We shall use the well-known Jacobian criterion. If we put $y_{00} = t/x_{22}$ and $y_{ij} = x_{ij}/x_{22}$ for (ij) different from (12), (13), (22), (23), the nine functions u_1, u_2, u_3, y_{ij} satisfy the following six identities

$$\begin{aligned} (y_{21})^2 + 1 - (u_3)^2 - (y_{00})^2 &= 0 \\ y_{11}y_{21} - iu_2 + iu_1u_3 &= 0 \\ (y_{11})^2 + (y_{21})^2 + (y_{31})^2 - (y_{00})^2 &= 0 \\ -(u_2)^2 + 1 + (y_{32})^2 - (y_{00})^2 &= 0 \\ -iu_1u_2 + iu_3 + y_{32}y_{33} &= 0 \\ -iu_2y_{33} - u_1y_{32} + y_{00}y_{21} &= 0. \end{aligned}$$

If we evaluate the Jacobian of this system with respect to the six y_{ij} at the image point of ∞, we get

$$\pm 2^3 (y_{21})^3 y_{31} (y_{32})^2 (\infty) = \pm 2^3 i,$$

and this is different from zero. Therefore u_1, u_2, u_3 form a set of local parameters of $\operatorname{proj}.A(\Gamma(2,4))$ at the image point of ∞.

Actually the same proof shows that the three functions in the lemma form a set of local co-ordinates of $\operatorname{proj}.A(\Gamma(2,4))$ at the image point of

$$\lim_{\eta\to\infty}\begin{pmatrix} w & 0 \\ 0 & i\eta \end{pmatrix}$$

for every w in the upper-half plane. In fact, the Jacobian in question will just become $\pm 2^8 i(\theta_{01}(w)/\theta_{00}(w))^6$.

LEMMA 5. *Let R be a* (noetherian) *normal domain and $a_1, a_2, \cdots, a_n$ be elements of R such that $Ra_i + Ra_j$ has rank two for $i \neq j$. Let b be a square root of the product $a_1 a_2 \cdots a_n$ and let S be a unique factorization domain which is integral over $R[b]$ and in which every unit has a square root. Then S contains the square roots of $a_1, a_2, \cdots, a_n$.*

Proof. Suppose that a square root b_1 of a_1, say, is not contained in S. Then in the decomposition of a_1 into irreducible elements in S, at least one of them, say P, appears odd times. Since the product $a_1 a_2 \cdots a_n$ is a square in S, another element, say a_2, also contains P odd times. Now the intersection $\mathfrak{p}$ of PS with R is a prime ideal of R, and it contains both a_1 and a_2. Therefore $\mathfrak{p}$ contains a minimal prime ideal $\mathfrak{q}$ of a_1, and $\mathfrak{q}$ is strictly smaller than $\mathfrak{p}$ by assumption. Hence the going-down theorem of Cohen-Seidenberg asserts that PS contains a prime ideal having $\mathfrak{q}$ as its intersection with R. But PS has rank one, and this is a contradiction.

Now, we shall start proving the theorem. We note that proj. $A(\Gamma(4))$ is an abelian covering of proj. $A(\Gamma(2,4))$ with the Galois group isomorphic to $\Gamma(2,4)/\Gamma(4)(\pm 1_4)$. Since this is a vector space over $\boldsymbol{Z}/2\boldsymbol{Z}$ of dimension five, we need five independent radicals, and the following elements of $A(\Gamma(4))$

$$\theta_{0000}\theta_{0001}\theta_{0010}\theta_{0011}, \qquad \theta_{0000}\theta_{0010}\theta_{0100}\theta_{0110}$$
$$\theta_{0000}\theta_{0001}\theta_{1000}\theta_{1001}, \qquad \theta_{0000}\theta_{0011}\theta_{1100}\theta_{1111}$$
$$\theta_{0000}\theta_{0100}\theta_{1000}\theta_{1100}$$

divided by $(\theta_{0000})^4$ are such radicals (as we see by TC, Theorem 3). We call them in this order $x_1, x_2, \cdots, x_5$. Then, up to units in the local ring of proj. $A(\Gamma(2,4))$ at the image point of ∞, we have

$$(x_1)^2 \sim 1, \qquad (x_2)^2 \sim u_1(u_2u_3 - u_1)$$
$$(x_3)^2 \sim u_2(u_3u_1 - u_2), \qquad (x_4)^2 \sim u_3(u_1u_2 - u_3)$$
$$(x_5)^2 \sim u_1u_2u_3.$$

Suppose therefore that a regular local ring S is integral over the analytic local ring of proj. $A(\Gamma(4))$ at the image point of ∞. Then S is isomorphic to a ring of convergent power-series in three variables over $\boldsymbol{C}$. In particular, it is a unique factorization domain in which every unit has a square root. Moreover Lemma 4 shows that the analytic local ring of proj. $A(\Gamma(2,4))$ at the image point of ∞ is the ring of convergent power-series in u_1, u_2, u_3

with coefficients in $\boldsymbol{C}$. Therefore we can apply Lemma 5 taking this regular local ring as R. We see, in this way, that S contains the square roots of u_1, u_2, u_3, $u_2u_3-u_1$, $u_3u_1-u_2$, $u_1u_2-u_3$. Namely, if we denote the square roots of u_1, u_2, u_3 by v_1, v_2, v_3, the ring S contains not only v_1, v_2, v_3 but also the square roots of $(v_1)^2-(v_2v_3)^2$, $(v_2)^2-(v_3v_1)^2$, $(v_3)^2-(v_1v_2)^2$. We apply Lemma 5 again to the same S and to the ring of convergent power-series in v_1, v_2, v_3 with coefficients in $\boldsymbol{C}$, and we see that

$$(v_1 \pm v_2v_3)^{\frac{1}{2}}, \qquad (v_2 \pm v_3v_1)^{\frac{1}{2}}, \qquad (v_3 \pm v_1v_2)^{\frac{1}{2}}$$

are all contained in S. We shall show that, if we put

$$\begin{aligned} u_1' &= 2(v_1+v_2v_3)/(1+u_1+u_2+u_3) \\ u_2' &= 2(v_2+v_3v_1)/(1+u_1+u_2+u_3) \\ u_3' &= 2(v_3+v_1v_2)/(1+u_1+u_2+u_3) \end{aligned}$$

and define v_1', v_2', v_3', u_1'', u_2'', u_3'', v_1'', v_2'', v_3'', $\cdots$ similarly, these two sequences are both contained in S.

We recall that, in TC, Section 4, we proved the following identity

$$\theta_{\mathfrak{m}}(\tau/2)^2 = \sum_{n' \bmod 2} (-1)^{{}^t m'' n'} \theta\binom{m'+n'}{0}(\tau)\,\theta\binom{n'}{0}(\tau).$$

In fact this was obtained by putting $z=0$ in the "second principal transformation" of degree two. At any rate, written explicitly, we have the following ten identities:

$$\begin{aligned}
\theta_{0000}(\tau/2)^2 &= \theta_{0000}(\tau)^2+\theta_{0100}(\tau)^2+\theta_{1000}(\tau)^2+\theta_{1100}(\tau)^2 \\
\theta_{0001}(\tau/2)^2 &= \theta_{0000}(\tau)^2-\theta_{0100}(\tau)^2+\theta_{1000}(\tau)^2-\theta_{1100}(\tau)^2 \\
\theta_{0010}(\tau/2)^2 &= \theta_{0000}(\tau)^2+\theta_{0100}(\tau)^2-\theta_{1000}(\tau)^2-\theta_{1100}(\tau)^2 \\
\theta_{0011}(\tau/2)^2 &= \theta_{0000}(\tau)^2-\theta_{0100}(\tau)^2-\theta_{1000}(\tau)^2+\theta_{1100}(\tau)^2 \\
\theta_{0100}(\tau/2)^2 &= 2(\theta_{0000}\theta_{0100}+\theta_{1000}\theta_{1100})(\tau) \\
\theta_{0110}(\tau/2)^2 &= 2(\theta_{0000}\theta_{0100}-\theta_{1000}\theta_{1100})(\tau) \\
\theta_{1000}(\tau/2)^2 &= 2(\theta_{0000}\theta_{1000}+\theta_{0100}\theta_{1100})(\tau) \\
\theta_{1001}(\tau/2)^2 &= 2(\theta_{0000}\theta_{1000}-\theta_{0100}\theta_{1100})(\tau) \\
\theta_{1100}(\tau/2)^2 &= 2(\theta_{0000}\theta_{1100}+\theta_{0100}\theta_{1000})(\tau) \\
\theta_{1111}(\tau/2)^2 &= 2(\theta_{0000}\theta_{1100}-\theta_{0100}\theta_{1000})(\tau).
\end{aligned}$$

Therefore, if we recall that we have

$$v_1 = \theta_{0100}/\theta_{0000}, \qquad v_2 = \theta_{1000}/\theta_{0000}, \qquad v_3 = \theta_{1100}/\theta_{0000},$$

the ten functions $\tau \rightarrow \theta_m(\tau/2)/\theta_{0000}(\tau)$ are all contained in S. Since $\tau \rightarrow \theta_{0000}(\tau/2)/\theta_{0000}(\tau)$ is one of the four units among them, the modified functions $\tau \rightarrow (\theta_m/\theta_{0000})(\tau/2)$ are also contained in S. We now observe that our u_1', u_2', u_3' and u_1, u_2, u_3 are related simply as

$$u_i'(\tau) = u_i(\tau/2)$$

for $i = 1, 2, 3$. It is thus clear that the entire argument can be repeated starting from u_1', u_2', u_3' instead of u_1, u_2, u_3. In fact, we have only to replace τ by $\tau/2$. Therefore, by an induction on n we see that the two sequences $u_i^{(n)}$, $v_i^{(n)}$ defined in an obvious way for $n = 1, 2, \cdots$ are contained in S. However, since we have the following relation

$$u_i = (u_i^{(n)})^{2^n}/2^{2(2^n-1)} + \cdots$$

for $n = 1, 2, \cdots$, we see that u_1, u_2, u_3 are contained in the intersection of powers of the ideal of non-units of S. Hence we get $u_1 = u_2 = u_3 = 0$, but this is a contradiction. Theorem 4 is thus proved.

It is not without some interest to see how the iteration process looks like in the elliptic case. Clearly we have to take $u = (\theta_{10}/\theta_{00})^2$, $v = \theta_{10}/\theta_{00}$. As for u' defined by $u'(w) = u(w/2)$, the formula we copied from TC tells that we have

$$u' = 2v/(1+u).$$

This is the quotient of the geometric mean of 1, u by their arithmetic mean. In other words, the process is the one which is familiar in the Gauss theory of "arithmetic-geometric means."

Also it is of some interest to extract a purely algebraic statement from the above proof:

COROLLARY 1. *Let K be an algebraically closed field of characteristic different from 2 and let R be a local ring which is integral over the ring of formal power-series in three variables u_1, u_2, u_3 over K and which contains a square root of the product $u_1u_2u_3(u_1 - u_2u_3)(u_2 - u_3u_1)(u_3 - u_1u_2)$. Then there is no regular local ring S which is integral over R.*

In fact S is isomorphic to a ring of formal power-series in three variables over K and, by Lemma 5, it will contain the square roots v_1, v_2, v_3 of u_1, u_2, u_3 and $(v_1 \pm v_2v_3)^{\frac{1}{2}}$, $(v_2 \pm v_3v_1)^{\frac{1}{2}}$, $(v_3 \pm v_1v_2)^{\frac{1}{2}}$. We define u_1', u_2', u_3' as before. If we can show that S contains the square root of the product

$$u_1'u_2'u_3'(u_1' - u_2'u_3')(u_2' - u_3'u_1')(u_3' - u_1'u_2'),$$

the same induction as before can be applied. Clearly S contains the square roots of u_1', u_2', u_3'. By the obvious symmetry, we have only to show that S contains the square roots of $u_1'-u_2'u_3'$. This is a consequence of the following relation

$$u_1'-u_2'u_3'=2(v_1-v_2v_3)(1+u_1-u_2-u_3)\cdot(1+u_1+u_2+u_3)^{-2},$$

which can be verified easily. Incidentally, in the complex case (hence in general by the universality) this reflects the following relation of theta-constants

$$(\theta_{0100}\theta_{0110})(\tau/2)=2(\theta_{0010}\theta_{0110})(\tau).$$

We note also that, in case K has a valuation like $\boldsymbol{C}$, we can talk about rings of convergent power-series instead of the rings of formal power-series, and get a similar result. This result can be used as a lemma (rather than a corollary) to prove Theorem 4 and, in this way, we can simplify the previous proof.

We shall now derive a more general statement from Theorem 4. The group Γ is assumed, as always, to be commensurable with $\Gamma(1)$:

COROLLARY 2. *If* Γ *has no element of finite order* (other than ± the identity), *there is no regular local ring which is integral over the analytic local ring of* proj. $A(\Gamma)$ *at the image point of* ∞.

We know that the intersection Γ' of Γ and $\Gamma(4)$ has finite indices in Γ and $\Gamma(4)$. Therefore proj. $A(\Gamma')$ is a (finite) covering of both proj. $A(\Gamma)$ and proj. $A(\Gamma(4))$. Suppose that there exists a regular local ring S which is integral over the analytic local ring R of proj. $A(\Gamma)$ at the image point of ∞. Let R' be the analytic local ring of proj. $A(\Gamma')$ at the image point of ∞. We shall denote by U, V, V' the (irreducible) local analytic varieties associated with S, R, R'. Then U and V, V' can be considered as "sheets" of $\boldsymbol{C}^3$ at the origin and of proj. $A(\Gamma)$, proj. $A(\Gamma')$ at the image points of ∞. Also we have covering morphisms $f\colon U\rightarrow V$ and $f'\colon V'\rightarrow V$. We shall show that there exists a covering morphism $h\colon U\rightarrow V'$ satisfying $f=f'h$. Let Y, Y' be the complements in V, V' of the quotient varieties $\Gamma\backslash\mathfrak{S}$, $\Gamma'\backslash\mathfrak{S}$, and let X be the inverse image in U. Then X, Y, Y' are analytic subsets of U, V, V' of co-dimension two and f' is "unramified" over $V-Y$. We take a point p of U not in X and pick a point q' of V' lying over $f(p)$. Then q' is not in Y' and since $U-X$ is simply connected, if we make an analytic continuation throughout $U-X$, the correspondence $p\rightarrow q'$ defines a unique (analytic) morphism of $U-X$ to $V'-Y'$. This morphism can be completed to a morphism h of U to V'. Clearly h satisfies $f=f'h$ and it is a covering morphism. On the other hand, since proj. $A(\Gamma')$ is a covering of proj. $A(\Gamma(4))$, we see that S is integral over the analytic local ring of proj. $A(\Gamma(4))$ at the image point of ∞. This contradicts Theorem 4.

We note that the above proof can be made more formal using the "purity of branch loci" due to Zariski and Nagata [cf. 8, pp. 158-68]. At any rate, since we know in general that $\Gamma_g(n)$ has no element of finite order for $n \geqq 3$, Corollary 2 can be applied to the analytic local rings of proj. $A(\Gamma(n))$ at the image points of ∞ for $n \geqq 3$. In this way, Theorem 4 becomes a special case of the more general Corollary 2.

5. Concluding remark. As before, let Γ be commensurable with $\Gamma(1)$. We shall denote by $v(\Gamma)$ the volume of the quotient variety $\Gamma \backslash \mathfrak{S}$ with respect to the invariant measure of the homogeneous space $\mathfrak{S}=Sp(2,\boldsymbol{R})/U(2)$. We know that $v(\Gamma)$ is finite [10]. Moreover, if Γ contains another Γ' as a subgroup of finite index, the quotient $v(\Gamma')/v(\Gamma)$ is equal to the index of $\Gamma'(\pm 1_4)$ in $\Gamma(\pm 1_4)$ and also to the degree of the covering proj. $A(\Gamma')$ $\rightarrow$ proj. $A(\Gamma)$. We normalize the constant factor in $v(\Gamma)$ so that we get

$$v(\Gamma(1)) = 1/2^6 3^2 5.$$

Actually, this "intrinsic volume" can be defined in general and in the elliptic case we have $v(\Gamma(1)) = 1/2^2 3$.

THEOREM 5. *According as Γ may or does not contain elements of finite orders, we have*

$$\dim_C A(\Gamma)_k = v(\Gamma)\cdot(k-1)(k-2)(\tfrac{1}{3}\cdot k-\tfrac{1}{2}) + 0(k^2), 0(k).$$

In the second case, it may be necessary that k tends to infinity as multiples of a certain integer.

We shall not give a proof of this theorem but leave it as an exercise to the reader. The proof we have in mind depends on Leray's spectral sequence, coherency of direct images of coherent sheaves and on the comparison of characteristic functions of coherent sheaves on a normal projective variety which coincide on some Zariski open set and finally on the following formula:

$$\dim_C A(\Gamma(4,8))_k = \begin{cases} 2^{11}3^{-1}k^3 - 2^{10}3k^2 + 2^8 3^{-1}67k - 2^7 31 \\ 1, 55, 695, 3969 \end{cases}$$

in which the first line is used for $k \geqq 4$ and the second line is used for

$k=0,1,2,3$. This result itself can be derived easily from Theorem 1 using some part of the proof of Lemma 1.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] H. Braun, " Zur Theorie der Modulformen n-ten Grades," *Mathematische Annalen*, vol. 115 (1938), pp. 507-517.
[2] Séminaire H. Cartan, Fonctions automorphes, E. N. S. (1957-58).
[3] G. Frobenius, " Über die Charaktere der symmetrischen Gruppe," *Sitzungsber. Berlin. Akad. Wiss.*, 1900, pp. 516-534.
[4] J. Igusa, " On the arithmetic normality of the Grassmann variety," *Proceedings of the National Academy of Sciences*, vol. 40 (1954), pp. 309-313.
[5] ———, " On Siegel modular forms of genus two," *American Journal of Mathematics*, vol. 84 (1962), pp. 175-200.
[6] ———, " On the graded ring of theta-constants," to appear (cited TC).
[7] A. Krazer-W. Wirtinger, "Abelsche Funktionen und algemeine Thetafunktionen," *Ensykl. Math. Wiss.*, II B 7, pp. 604-873.
[8] M. Nagata, *Local rings*, New York (1962).
[9] C. L. Siegel, " Einführung in die Theorie der Modulfunktionen n-ten Grades," *Mathematische Annalen, vol.* 116 (1939), pp. 617-657.
[10] ———, " Symplectic geometry," *American Journal of Mathematics*, vol. 65 (1943), pp. 1-86.
[11] H. Weyl, *Classical groups*, Princeton (1946).
[12] O. Zariski, " The concept of a simple point of an abstract algebraic variety," *Transactions of the American Mathematical Society*, vol. 62 (1947), pp. 1-52.

FINITELY GENERATED KLEINIAN GROUPS.*

By Lars V. Ahlfors.

---

1. Introduction. 1.1. Let Γ be a group of linear transformations

$$Az = \frac{az+b}{cz+d}, \quad ad - bc = 1$$

of the extended complex plane. A point z_0 is called a *limit point* of Γ if it is an accumulation point of points Az_1, $A \in \Gamma$, some some z_1. The set of limit points will be denoted by $\Sigma(\Gamma)$, or by Σ when it is clear to what group we are referring.

$\Sigma(\Gamma)$ is a closed set, and invariant under Γ; we denote its complement by $\Omega = \Omega(\Gamma)$. As soon as $\Omega(\Gamma)$ is not empty we say that Γ is discontinuous, and $\Omega(\Gamma)$ is its *set of discontinuity*. It is easy to classify all groups for which Σ is void, consists of a single point, or of two points. All other discontinuous groups will be called *Kleinian groups*. In other words, a Kleinian group is a discontinuous group with more than two limit points. In this paper Γ will always denote a Kleinian group.

For Kleinian groups the images Az_1, $A \in \Gamma$, of any point z_1 accumulate toward all of Σ. It is an immediate consequence that Σ is a nowhere dense perfect set, and that every nonvoid closed invariant set includes Σ.

A Kleinian group which leaves a circle invariant is said to be Fuchsian. The set of limit points lies on the invariant circle. If Σ is the whole circle the Fuchsian group is of the first kind. If not, Σ is nowhere dense on the invariant circle, and the group is of the second kind.

The fixpoints of non-elliptic transformations in Γ lie on Σ. There are always infinitely many hyperbolic or loxodromic transformations, and their fixpoints are dense on Σ. The fixpoints in Ω belong to elliptic transformations, and they are isolated.

1.2. In order to study the action of Γ on Ω we consider the quotient space $S = \Omega/\Gamma$. It has a natural complex structure such that the projection map $\pi: \Omega \to S$ is holomorphic. Thus, the components S_i of S are Riemann

---

Received October 30, 1963.

* This work was supported (in part) by the Air Force Office of Scientific Research.

surfaces. We shall write $\Omega_i = \pi^{-1} S_i$. In general, the Ω_i are not connected, and we denote the components of Ω_i by Ω_{ij}.

With any Kleinian group Γ we have thus associated the decompositions $\Omega = \cup\, \Omega_i = \cup\, \Omega_{ij}$ and $S = \cup\, S_i$. Each Ω_i is invariant under the full group Γ, and the boundary of Ω_i is all of Σ. The components Ω_{ij} are ramified covering surfaces of S_i whose branch points are elliptic fixpoints.

The projections of the branch points will be denoted by p_k. They are isolated, and with each p_k there is associated an integer $m_k \geqq 2$, the order of the corresponding elliptic transformation. We say that S has a signature $I = \{p_k, m_k\}$. S together with its signature is denoted by $S(I)$, and $S_i(I)$ has a similar meaning.

We shall say that S is of *finite type* if there are finitely many components S_i and points p_k, and if each S_i can be obtained from a compact surface by omission of a finite number of points. When this is so it is convenient to let S_i denote the compact surface, and to regard the omitted points as points p_k with $m_k = \infty$. It will be clear from the context when S_i should be interpreted in this manner.

1.3. The main result to be proved in this article is the following statement:

If Γ *is finitely generated, then* S *is of finite type.*

This is known for Fuchsian groups, but to the author's knowledge there is no proof in the published literature. A proof is contained in a manuscript of Nielsen and Fenchel, where the result is obtained at the end of a penetrating study of Fuchsian groups. Recently, L. Bers has found a short proof by variational methods, the same that have proved so useful in the theory of Teichmüller spaces.[1]

My task has been to show that Bers' method can be extended to arbitrary Kleinian groups. The main difficulty was the correct formulation and proof of a crucial lemma, which is to be considered the core of this paper. We have formulated it in two parts, as Lemma 8 and Lemma 9.

In the last section we prove some results on functions groups, that is, groups with an invariant region. These were the results that gave the initial impetus to this study.

1.4. Perhaps of greater interest are the theorems I have not been able to prove. For instance, it seems very likely that the set Σ for any finitely generated group Γ has zero area, but I cannot prove it.

---

[1] Unpublished.

Any Kleinian group can be extended to a corresponding group of Möbius transformations in the Poincaré half-space. If Γ is finitely generated it seems plausible that the Poincaré group has a fundamental polyhedron with only a finite number of vertices on Σ. Again, I cannot prove this, but if this is so I am able to show that Σ has indeed zero measure.

2. The Poincaré metric. 2.1. Each component Ω_{ij} carries a Poincaré metric $ds = \rho\,|dz|$ with constant negative curvature equal to -1, and such that Ω_{ij} is complete in this metric. The metric is unique, and therefore

$$\rho(Az)\,|A'(z)| = \rho(z) \tag{2.1}$$

for all $A \in \Gamma$ which map Ω_{ij} onto itself. We shall consider ρ on all of Ω, defined in each Ω_{ij}. Then (2.1) holds for all $A \in \Gamma$.

The projection π induces a corresponding metric $\bar\rho\,|d\zeta|$ on S. Here ζ is a local parameter on S, and $\bar\rho$ is determined by $\bar\rho\,|d\zeta| = \rho|\,dz\,|$. At the branch points $d\zeta/dz = 0$, and we see that $\bar\rho$ becomes infinite at the points p_k.

If S_i is compact the Gauss-Bonnet formula yields

$$\chi(S_i) + \sum_{p_k \in S_i} \Big(1 - \frac{1}{m_k}\Big) = \frac{1}{2\pi}\int_{\Omega_i/\Gamma} \rho^2\,dxdy, \tag{2.2}$$

where $\chi(S_i)$ is the Euler characteristic. The formula remains valid for any $S_i(I)$ of finite type, provided that $\chi(S_i)$ is the characteristic of the compact surface and the infinite signatures are included in the sum. To prove (2.2) for this case requires knowledge of the asymptotic behavior of $\bar\rho$ at the punctures. If the omitted point corresponds to $\zeta = 0$ this behavior is given by

$$\log\bar\rho = -\log|\zeta| - \log\log(1/|\zeta|) + O(1), \tag{2.3}$$

a classical result.

An immediate consequence of (2.2) is that

$$\chi(S_i) + \sum_{p_k \in S_i} \Big(1 - \frac{1}{m_k}\Big) > 0. \tag{2.4}$$

This rules out certain possibilities. Without insisting on the details, we conclude that S_i cannot be a sphere with certain low signatures, nor can it be a torus without signature.

A more careful analysis of the possible cases shows that (2.4) can be sharpened to

$$\chi(S_i) + \sum_{S_i} \Big(1 - \frac{1}{m_k}\Big) \geqq 1/42.$$

This is important because it shows that S *is of finite type if and only if its* *total Poincaré area is finite.* The conclusion uses the elementary fact that a single surface with finite Poincaré area is necessarily of finite type.

2.2. We show next that the points p_k with $m_k = \infty$ correspond to parabolic fixpoints. Since the result does not require S to be of finite type, let us merely assume that $S_i = \hat{S}_i - \{p_k\}$ where $\hat{S}_i$ is a Riemann surface and $p_k \in \hat{S}_i$.

LEMMA 1. *There exists a parabolic transformation* $A_k \in \Gamma$ *with the following property: If the linear transformation* U *is such that* $U^{-1}A_kUz = z+1$, *then* Ω_i *contains a set* $\{z \mid \operatorname{Im} Uz > c\}$ *and* $\pi(z) \to p_k$ *when* $z \in \Omega_i$ *approaches the fixpoint* $z_k = U\infty$ *in such a way that* $\operatorname{Im} Uz \to \infty$.

We remark that A_k can of course be replaced by any conjugate transformation $B^{-1}A_kB$ with $B \in \Gamma$, in which case z_k is replaced by Bz_k. The set $\{z \mid \operatorname{Im} Uz > 0\}$ is a disk or a halfplane whose boundary passes through z_k. It follows that there are at most two p_k which correspond to the same z_k.

For the proof of the lemma we may regard p_k as the center of a punctured disk $\tilde{\Delta} \subset S_i$, represented by $0 < |\zeta| < 1$, and we may assume that no point in $\tilde{\Delta}$ has a signature. Let Δ be a component of $\pi^{-1}(\tilde{\Delta})$. It is clear that the cover transformations of Δ over $\tilde{\Delta}$ are generated by a single transformation $A \in \Gamma$ which is not elliptic.

If A is hyperbolic or loxodromic we can take it to be of the form $Az = kz$, $|k| \neq 1$. In terms of $s = \log \zeta$ the projection π induces a one-one correspondence $z = z(s)$ between the halfplane $\operatorname{Re} s < 0$ and Δ which satisfies $z(s+2\pi i) = kz(s)$. Because 0 is a limit point it is not in Δ, and therefore the distance from $z(s)$ to the boundary of Δ is at least $|z(s)|$. With this information Koebe's one-quarter theorem easily yields

$$|z'(s)|/|z(s)| \leqq 2/|\operatorname{Re} s|.$$

Integration from s to $s+2\pi i$ along a vertical line gives

$$|\log |k|| \leqq 4\pi/|\operatorname{Re} s|.$$

But this is absurd, for $|\operatorname{Re} s|$ can be chosen arbitrarily large.

We conclude that A is parabolic, for instance $Az = z+1$. In this case $z(s+2\pi i) = z(s)+1$ and it is elementary to deduce that

$$\exp(2\pi i z) = \zeta + a_0 + a_1/\zeta + \cdots,$$

from which

$$z(s) = \frac{s}{2\pi i} + O(|\zeta|^{-1}) \tag{2.5}$$

and, for later use,

$$(2.6)\qquad dz/d\zeta = (2\pi i)^{-1} + O(|\zeta|^{-2}).$$

It is immediate from (2.5) that Ω_i contains a halfplane $y > c$, and that $y \to \infty$ implies $\pi(z) \to p_k$. This is the assertion of the lemma when U is the identity, and the general case follows by application to the group $U^{-1}\Gamma U$.

2.3. We shall use (2.6) to derive the asymptotic behavior of ρ in terms of the uniformizer ζ. Because $\rho\,|dz| = \bar\rho\,|d\zeta|$ we obtain, with the help of (2.3),

$$(2.7)\qquad \log\rho = -\log\log(1/|\zeta|) + O(1).$$

In the derivation of this formula we have assumed that U is the identity, but it is easy to see that it remains valid as soon as the fixpoint is at ∞.

In case of a finite fixpoint we make the transformation $z = Uw$. The Poincaré metric in $U^{-1}\Omega$ is given by $\rho_0(w)\,|dw| = \rho(z)\,|dz|$. It is for ρ_0 that (2.7) is valid, while (2.5) is applicable to w in the place of z. We have $\log\rho = \log\rho_0 - \log|U'(w)| = \log\rho_0 + 2\log|w| + O(1)$, for since $U\infty$ is finite, $w^2U'(w)$ has a finite limit when $w \to \infty$. It follows that

$$\log\rho = -\log\log(1/|\zeta|) + 2\log|\log\zeta| + O(1).$$

The expression on the right is not single-valued, but we may conclude that

$$(2.8)\qquad \log\rho = \log\log(1/|\zeta|) + O(1)$$

when z approaches the fixpoint in an interior angle.

2.4. The coefficient of the Poincaré metric decreases when the region increases (this is essentially Schwarz' lemma). Denote the euclidean distance from z to Σ by $\delta(z)$. If we compare Ω with the circle about z of radius $\delta(z)$ it follows at once that $\rho(z) \leqq 2\delta(z)^{-1}$. In case Ω_{ij} is simply connected and does not contain ∞ an opposite inequality can be obtained by use of Koebe's one-quarter theorem. One finds under these circumstances $\rho(z) \geqq \frac{1}{2}\delta(z)^{-1}$.

In the general case we use the majorization $\Omega \subset \Omega_0$, where Ω_0 is the sphere punctured at three points $z_0, z_1, z_2 \in \Sigma$. We have $\rho \geqq \rho_0$, and ρ_0 can be estimated by means of (2.3). It follows that

$$\log\rho(z) \geqq -\log|z - z_0| - \log\log(1/|z - z_0|) - O(1)$$

in a neighborhood of z_0. It is not hard to see that the remainder is uniformly bounded when z_0 varies but stays away from z_1, z_2. If Σ is assumed to be compact we conclude that

$$(2.9)\qquad \log\rho \geqq -\log\delta - \log\log(1/\delta) - O(1)$$

for sufficiently small δ.

3. Quadratic differentials. 3.1. In all that follows a dominant role is played by the quadratic differentials on Ω. A quadratic differential is a holomorphic function ϕ on Ω which satisfies the functional equations

$$\phi(Az)A'(z)^2 = \phi(z) \tag{3.1}$$

for all $A \in \Gamma$. More accurately, the quadratic differential is the invariant expression $\phi\, dz^2$, but it is convenient to use the same term for the coefficient ϕ.

The projection map induces a corresponding quadratic differential on S, namely by the relation $\phi\, dz^2 = \bar{\phi}\, d\zeta^2$. At a point p_k with finite signature, let the projection be given by $\zeta = z^m$ in local coordinates. Then $\bar{\phi} = O(|\zeta|^{-2+2/m})$, and since $\bar{\phi}$ is single-valued it has at most a simple pole at p_k. Conversely, if $\bar{\phi}$ has at most a simple pole, then ϕ is holomorphic.

As soon as S is not compact the class of all quadratic differentials on Ω is too extensive to be useful. We shall consider the following restriction:

$$\int_{\Omega/\Gamma} |\phi|\, dxdy < \infty, \tag{Q}$$

$$|\phi| = O(\rho^2). \tag{Q^*}$$

Observe that the integral in (Q) is meaningful because of (3.1). The quadratic differentials which satisfy these conditions form linear spaces which we denote by $Q(\Gamma)$ and $Q^*(\Gamma)$ respectively.

In terms of $\bar{\phi}$ condition (Q) becomes

$$\int_S |\bar{\phi}|\, d\xi d\eta < \infty. \tag{3.2}$$

Similarly, (Q^*) yields $|\bar{\phi}| = O(\bar{\rho}^2)$. In both cases $\bar{\phi}$ has at most a simple pole at points with infinite signature. In fact, for Q this follows directly from (3.2), and for Q^* we may use (2.3) to deduce that $\bar{\phi} = O(|\zeta|^{-2})$ in terms of a local variable, and hence that the singularity is at most a simple pole.

We shall let $\bar{Q}(S)$ and $\bar{Q}^*(S)$ denote the spaces formed by all $\bar{\phi}$ for $\phi \in Q(\Gamma)$ and $\phi \in Q^*(\Gamma)$ respectively. If S is of finite type these spaces are identical, and we conclude:

LEMMA 2. *If S is of finite type the spaces $Q(\Gamma)$ and $Q^*(\Gamma)$ are identical and of finite dimension.*

The dimension over the complex numbers is $3g + n - 3k$ where g is the sum of the genera of the surfaces S_i, n is the total number of points with a signature, and k is the number of components. It follows from condition (2.4) that there are non-zero quadratic differentials on each S_i.

3.2. We have just proved one part of the following result, which is particularly important for our purposes:

THEOREM 1. *The space* $Q(\Gamma)$ *is finite dimensional if and only if* S *is of finite type.*

If there are infinitely many S_i, all of finite type, it follows from the remark above that $Q(\Gamma)$ has infinite dimension. Suppose now that one S_i is not of finite type. If S_i has infinite genus there exist infinitely many linearly independent square integrable first order differentials θ_j on S_i. If they are multiplied by a fixed one, θ_0, the $\theta_0\theta_j$ are linearly independent quadratic differentials, and they are integrable. Assume next that S_i has finite genus. Then it can be imbedded in a compact Riemann surface $\hat{S}_i$. If $\hat{S}_i - S_i$ is an infinite point set we can find infinitely many quadratic differentials on $\hat{S}_i$ with distinct simple poles on $\hat{S}_i - S_i$. Their restrictions to S_i are integrable and linearly independent. Finally, if $\hat{S}_i - S_i$ is a finite set there must be infinitely many p_k on S_i, and we can find infinitely many quadratic differentials with distinct poles among the p_k. Thus, whenever S is not of finite type, $Q(\Gamma)$ has infinite dimension.

4. Beltrami differentials. 4.1. We regard $Q(\Gamma)$ as a Banach space with the norm

$$\| \phi \| = \int_{\Omega/\Gamma} | \phi | \, dxdy$$

and $Q^*(\Gamma)$ as one with the norm

$$\| \phi \|_* = \sup | \phi | \rho^{-2}.$$

One of our aims is to determine the conjugate space of $Q(\Gamma)$. Several characterizations will be obtained, and their comparison will contribute to our knowledge of Kleinian groups.

4.2. Choose an arbitrary $\phi_0 \in Q(\Gamma)$, not identically zero, and define a measure on $S = \Omega/\Gamma$ by $dm = | \phi_0 | \, dxdy$. Evidently, a quadratic differential ϕ belongs to $Q(\Gamma)$ if and only if ϕ/ϕ_0 is of class L^1 with respect to this measure. Therefore, every linear functional on $Q(\Gamma)$ can be represented as

$$\int_{\Omega/\Gamma} \phi_0^{-1} | \phi_0 | \beta \, dxdy$$

where β is a bounded measurable function on S. We may define β as an automorphic function on Ω, and we find that $\nu = \phi_0^{-1} | \phi_0 | \beta$ satisfies the equations

$$\nu(Az)\bar{A}'(z)/A'(z) = \nu(z) \tag{4.1}$$

for all $A \in \Gamma$. Such functions are called *Beltrami differentials*; we emphasize that ν has to be bounded and measurable. We have proved that every linear functional on $Q(\Gamma)$ is of the form

$$\int_{\Omega/\Gamma} \phi\nu \, dxdy \tag{4.2}$$

where ν is a Beltrami differential.

4.3. Let $B(\Gamma)$ be the linear space of Beltrami differentials, normed by $\| \nu \|_\infty$, and let $N(\Gamma)$ denote the subspace of those ν which satisfy

$$\int_{\Omega/\Gamma} \phi\nu \, dxdy = 0 \tag{4.3}$$

for all $\phi \in Q(\Gamma)$. Our result can be stated as follows:

LEMMA 3. *The conjugate space of* $Q(\Gamma)$ *can be identified with* $B(\Gamma)/N(\Gamma)$.

Clearly, any characterization of $N(\Gamma)$ will therefore be a characterization of the conjugate space of $Q(\Gamma)$.

5. The Fuchsian case. 5.1. Many problems in the theory of Kleinian groups can be reduced to the corresponding problem for Fuchsian groups. We shall therefore begin with a study of the Fuchsian case.

To emphasize the distinction, Fuchsian groups will be denoted by the letter G. We assume that G is discontinuous and acts on the unit disk $D = \{z \mid |z| < 1\}$, but it need not be of the first kind. We shall constantly disregard the outside of D. Therefore, we shall set $S(G) = D/G$, and the notations $Q(G)$ and $Q^*(G)$ will refer to quadratic differentials in D. The corresponding remark applies to the spaces $B(G)$ and $N(G)$. Observe that $\rho(z) = 2(1 - |z|^2)^{-1}$.

We shall make frequent use of the known representation formula[2]

$$\phi(\zeta) = 3\pi^{-1} \int_D \phi(z)(1 - |z|^2)^2 (1 - \bar{z}\zeta)^{-4} \, dxdy \tag{5.1}$$

---

[2] L. Ahlfors, "Some remarks on Teichmüller's space of Riemann surfaces," *Annals of Mathematics*, vol. 74, No. 1, 1961, p. 176.

which is valid as soon as ϕ is holomorphic in D and

$$\int_D |\phi|(1-|z|^2)^2\, dxdy < \infty. \tag{5.2}$$

This condition is obviously fulfilled when $\phi \in Q^*(G)$.

The identity

$$\int_D |1-\bar{z}\zeta|^{-4}\, dxdy = \pi(1-|\zeta|^2)^{-2} \tag{5.3}$$

will also be needed.

5.2. With any Beltrami differential $\nu \in B(G)$ we associate

$$T\nu(\zeta) = 3\pi^{-1}\int_D \bar{\nu}(z)\,(1-\bar{z}\zeta)^{-4}\, dxdy. \tag{5.4}$$

A simple computation shows that $T\nu$ is a quadratic differential, and by use of (5.3) we conclude that $T\nu \in Q^*(G)$. In the opposite direction, $\phi \in Q^*(G)$ can be mapped on $T^*\phi = \bar{\phi}(1-|z|^2)^2 = 4\bar{\phi}\rho^{-2} \in B(G)$. Formula (5.1) takes the form $TT^*\phi = \phi$, or $TT^* = I$. The operator T^* is of course meaningful also when $\phi \in Q(G)$, but then we can no longer assert that $T^*\phi$ belongs to $B(G)$.

As an abbreviation we introduce

$$\langle\phi, \nu\rangle = \int_{D/G} \phi\nu\, dxdy, \tag{5.5}$$

and we use this notation as soon as the integral on the right is absolutely convergent. One verifies at once that

$$\langle\phi_1, T^*\phi_2\rangle = \langle\phi_2, T^*\phi_1\rangle^- \tag{5.6}$$

as soon as $\phi_1 \in Q(G)$, $\phi_2 \in Q^*(G)$ or vice versa (the bar stands for the complex conjugate).

Even more important is the following identity:

LEMMA 4. *$\langle\phi, T^*T\nu\rangle = \langle\phi, \nu\rangle$ for $\phi \in Q(G)$, $\nu \in B(G)$.*

Both sides of the identity represent absolutely convergent integrals, and by (5.6) the left hand side can be replaced by $\langle T\nu, T^*\phi\rangle^-$. By virtue of (5.1) the equality $\langle T\nu, T^*\phi\rangle^- = \langle\phi, \nu\rangle$ is equivalent to

$$\int_{D/G} \phi(\zeta)\,(1-|\zeta|^2)^2\, d\xi d\eta \int_D \nu(z)\,(1-\bar{z}\zeta)^{-4}\, dxdy$$
$$= \int_D \phi(\zeta)\,(1-|\zeta|^2)^2\, d\xi d\eta \int_{D/G} \nu(z)\,(1-\bar{z}\zeta)^{-4}\, dxdy,$$

an identity which is proved by subdivision into "fundamental regions", and subsequent changes of the integration variables. The absolute convergence justifies the procedure.

5.3. Let $A(D)$ be the space of holomorphic functions which are integrable over D. For $F \in A(D)$ the Poincaré series

$$\Theta F(z) = \sum_{A \in G} F(Az) A'(z)^2 \tag{5.7}$$

converges and represents an element of $Q(G)$. In fact, one obtains

$$\| \Theta F \| \leqq \int_D |F| \, dxdy. \tag{5.8}$$

It is also easy to verify that

$$\langle \Theta F, \nu \rangle = \int_D F \nu \, dxdy. \tag{5.9}$$

5.4. We now deduce the most important result in this direction, which is due to L. Bers:[3]

THEOREM 2. *The mapping* $T: B(G) \to Q^*(G)$ *is a surjection with kernel* $N(G)$. *Therefore, the conjugate space of* $Q(G)$ *can be identified with* $Q^*(G)$.

That the mapping is onto follows from $TT^* = I$. Another way of expressing Lemma 4 is to say that $\nu - T^*T\nu \in N(G)$. Therefore, $T\nu = 0$ implies $\nu \in N(G)$. To prove the converse, let F be the function $3\pi^{-1}(1 - z\bar{\zeta})^{-4}$ with fixed $\zeta \in D$. The definition (5.4) of $T\nu$ together with (5.9) gives

$$T\nu(\zeta)^- = \langle \Theta F, \nu \rangle,$$

and hence $\nu \in N(G)$ yields $T\nu = 0$ as asserted.

Specifically, we have shown that every coset of $N(G)$ in $B(G)$ contains an element $T^*\phi$ with $\phi \in Q^*(G)$, namely for $\phi = T\nu$. This is of course the only element with this property, for $T^*\phi \in N(G)$ implies $\phi = TT^*\phi = 0$.

5.5. As an immediate corollary we obtain:

LEMMA 5. *Every* $\phi \in Q(G)$ *can be written in the form* ΘF *with* $F \in A(D)$.

Indeed, it follows from (5.8) that Θ maps $A(D)$ onto a closed subspace of $Q(G)$. If it were not all of $Q(G)$ we could find $\phi \in Q^*(G)$, not identically zero, such that $\langle \Theta F, T^*\phi \rangle = 0$ for all $F \in A(D)$. By (5.9) this gives

---

[3] Oral communication.

$$\int_D FT^*\phi\, dxdy = 0.$$

We choose again $F(z) = 3\pi^{-1}(1-z\bar{\zeta})^{-4}$ and obtain $\phi(\zeta) = 0$, contrary to assumption.

5.6. There is a generalization of Lemma 5 which will be of importance when we pass to the case of Kleinian groups. Let G_0 be an arbitrary subgroup of G. For any $\phi_0 \in Q(G)$ we construct

$$\Theta_0\phi_0(z) = \sum \phi_0(Az)A'(z)^2 \tag{5.10}$$

where the summation is now over a set of representatives of the left cosets of G_0, one from each coset. Because ϕ_0 is a quadratic differential with respect to G_0 the choice of representatives is irrelevant, and one finds that $\| \Theta_0\phi_0 \| \leqq \| \phi_0 \|$. Equation (5.9) is replaced by

$$\langle \Theta_0\phi_0, \nu\rangle = \langle \phi_0, \nu\rangle, \tag{5.11}$$

valid for all $\nu \in B(G)$. It is to be observed that the first inner product is over D/G, the second over D/G_0.

Lemma 6. *Every $\phi \in Q(G)$ can be written as $\Theta_0\phi_0$ with $\phi_0 \in Q(G_0)$.*

Again, if this were not true we could find $\phi \in Q^*(G)$, $\phi \neq 0$, such that $\langle \Theta_0\phi_0, T^*\phi\rangle = 0$ for all $\phi_0 \in Q(G)$. According to (5.11) this gives $\langle \phi_0, T^*\phi\rangle = 0$ for all ϕ_0, and hence $\phi = 0$, a contradiction.

6. The general case. 6.1. We return to the case of an arbitrary Kleinian group Γ, except that we shall assume ∞ to be a limit point. This minor restriction has the advantage that the Poincaré series

$$\Theta F(z) = \sum_{A\in\Gamma} F(Az)A'(z)^2 \tag{6.1}$$

remains holomorphic for any $F \in A(\Omega)$, that is, whenever F is holomorphic and integrable in Ω. It is clear that ΘF enjoys the same properties as in the Fuchsian case.

We shall show that the conclusion in Theorem 2 as well as Lemma 5 remain valid. As far as the conjugate space is concerned it is quite clear that we may restrict our attention to a single component Ω_{ij} of Ω and replace Γ by the subgroup Γ_{ij} which leaves Ω_{ij} invariant. It amounts to the same thing if we assume from the beginning that Ω is an invariant region with respect to Γ, but not necessarily the whole set of discontinuity.

Let $z = \omega(\zeta)$ be a conformal mapping of $D = \{\zeta \mid |\zeta| < 1\}$ onto the universal covering surface of Ω. We introduce a Fuchsian group G as follows: $B \in G$ if and only if there exists $A \in \Gamma$ such that

$$\omega(B\zeta) = A\omega(\zeta). \tag{6.2}$$

It is evident that A is uniquely determined, and the mapping $A \to B$ is a surjective homomorphism whose kernel we denote by G_0.

The equation

$$\hat{\phi}(\zeta) = \phi(\omega(\zeta))\omega'(\zeta)^2 \tag{6.3}$$

sets up a one-one correspondence between quadratic differentials with respect to G and Γ. One verifies that $Q(\Gamma)$ corresponds to $Q(G)$ and $Q^*(\Gamma)$ to $Q^*(G)$. A similar correspondence between $B(\Gamma)$ and $B(G)$ is defined by

$$\hat{\nu}(\zeta) = \nu(\omega(\zeta))\overline{\omega}'(\zeta)/\omega'(\zeta). \tag{6.4}$$

The identity

$$\int_{\Omega/\Gamma} \phi\nu \, dxdy = \int_{D/G} \hat{\phi}\hat{\nu} \, d\xi d\eta \tag{6.5}$$

prevails. We conclude:

THEOREM 3. *The conjugate space of $Q(\Gamma)$ has a canonical identification with $Q^*(\Gamma)$.*

6.2. For the counterpart of Lemma 5 it is also true that it suffices to consider a component Ω_{ij} and the corresponding subgroup Γ_{ij}. To see this, let Θ_{ij} denote the theta-operator with respect to Γ_{ij}. If $\phi = \Theta_{ij}F$ in Ω_{ij} with $F \in A(\Omega_{ij})$, extend F to all of Ω by setting $F = 0$ outside of Θ_{ij}. At the same time ϕ is extended to all of Ω_i as a quadratic differential. It follows that $\phi = \Theta F$ in Ω_i, and this is what we wish to prove. It is therefore no restriction to assume again that Ω is an invariant region for Γ.

LEMMA 7. *Every $\phi \in Q(G)$ can be written as ΘF with $F \in A(\Omega)$.*

We use the same notations as before. By Lemma 6 it is possible to write $\phi = \Theta_0\phi_0$ with $\phi_0 \in Q(G_0)$. The relation

$$\phi_0(\zeta) = F(\omega(\zeta))\omega'(\zeta)^2 \tag{6.6}$$

defines F uniquely, and one verifies that $\Theta_0\phi_0(\zeta) = \Theta F(\omega(\zeta))\omega'(\zeta)^2$, and this makes $\phi = \Theta F$.

7. Direct characterization of $N(\Gamma)$. 7.1. In this section we shall derive some results which, as far as we are able to judge, cannot be obtained by reduction to the Fuchsian case. It will be convenient to assume, in this connection, that the points 0, 1, ∞ belong to Σ. All invariantly formulated results are of course independent of this hypothesis.

The method we shall use is intimately connected with infinitesimal deformations of conformal structure. However, there will be no need to make this connection explicit.

Given $\nu \in B(\Gamma)$ we construct

$$f(\zeta) = -\pi^{-1} \int_{\Omega} \nu(z) \left[\frac{1}{z-\zeta} - \frac{\zeta}{z-1} - \frac{1-\zeta}{z}\right] dxdy. \tag{7.1}$$

The integral converges because the rational function in brackets has a third order zero at ∞ and only simple poles at ζ, 0, and 1. It represents a continuous function in the finite plane. What is more, it satisfies a uniform Hölder-type condition

$$|f(\zeta_1) - f(\zeta_2)| \leqq C\,|\zeta_1 - \zeta_2|\log 1/|\zeta_1 - \zeta_2|, \tag{7.2}$$

for instance for $|\zeta_1 - \zeta_2| \leqq 1/2$. At ∞ it is of order $O(|\zeta|\log|\zeta|)$.

It is well known that f has locally square integrable distributional derivatives, and that

$$f_{\bar z} = \tfrac{1}{2}(f_x - i f_y) = \nu \tag{7.3}$$

almost everywhere (provided that we set $\nu = 0$ on Σ).

7.2. We shall prove:

Lemma 8. *$\nu \in N(\Gamma)$ if and only if $f = 0$ on Σ.*

The necessity is immediate. Indeed, for any fixed $\zeta \in \Sigma$ the bracketed expression in (7.1) is a holomorphic and integrable function F on Ω. Therefore, $\nu \in N(\Gamma)$ implies

$$\int_{\Omega} \nu F\, dxdy = \langle \Theta F, \nu \rangle = 0,$$

and hence $f(\zeta) = 0$.

To prove the sufficiency, let λ be a C^∞ function which vanishes in a neighborhood of Σ. In addition we require that λ and $f\lambda_{\bar z}$ be bounded. Under these conditions, if $F \in A(\Omega)$ Stokes' formula yields, with the help of (7.3),

$$\int_{\Omega} \lambda F \nu\, dxdy = \int_{\Omega} f_{\bar z} \lambda F\, dxdy = -\int_{\Omega} f \lambda_{\bar z} F\, dxdy. \tag{7.4}$$

If we can choose λ, depending on a parameter, in such a way that λ tends boundedly to 1 and $f\lambda_{\bar z}$ tends boundedly to 0, it will follow that $\langle \Theta F, \nu \rangle = 0$, and hence, by Lemma 7, that $\nu \in N(\Gamma)$.

We are assuming that $f = 0$ on Σ. As before, $\delta(z)$ denotes the distance from z to Σ. We deduce from (7.2) that

$$|f| \leqq C\delta \log 1/\delta \tag{7.5}$$

for $\delta \leqq \delta_0$, say. We choose $\delta_0 \leqq e^{-1}$.

For the moment, let us ignore the condition that λ be of class C^∞. Given $\epsilon > 0$, let $\Lambda(t)$ be 0 for $0 \leqq t \leqq \epsilon$, 1 for $t \geqq 2\epsilon$, and linear between ϵ and 2ϵ. We shall set

$$\lambda(z) = \Lambda[(\log\log 1/\delta(z))^{-1}] \tag{7.6}$$

if $\delta < \delta_0$ and $\lambda(z) = 1$ if $\delta \geqq \delta_0$. For ϵ small enough, λ will be continuous and identically zero near Σ.

If $\delta(z)$ were differentiable, the obvious inequality $|\delta(z_1) - \delta(z_2)| \leqq |z_1 - z_2|$ would yield $|\delta_{\bar{z}}| \leqq 1$, and hence

$$|\lambda_{\bar{z}}| \leqq \epsilon^{-1}\delta^{-1}(\log 1/\delta)^{-1}(\log\log 1/\delta)^{-2}$$

in the part where λ is not constant. Since $(\log\log 1/\delta)^{-1} \leqq 2\epsilon$ for these values it follows that

$$\psi\,|\lambda_{\bar{z}}| \leqq 4\epsilon\delta^{-1}(\log 1/\delta)^{-1},$$

and in view of (7.5) we obtain $|\lambda_{\bar{z}} f| \leqq 4C\epsilon$ throughout Ω. The desired conclusion follows.

The reasoning is completed by an obvious smoothing process, which we need not explain in detail. First, we replace δ in (7.6) by a smooth mean value δ^* obtained by convolution. In the relevant part of the plane δ^* will satisfy the same Lipschitz condition as δ. Secondly, Λ has to be replaced by a smooth function whose derivative is nowhere much greater than ϵ^{-1}. With these modifications the argument goes through exactly as before.

7.3. An alternative characterization of $N(\Gamma)$ is the following:

LEMMA 9. *$\nu \in N(\Gamma)$ if and only if $f(Az) = f(z)A'(z)$ for all $A \in \Gamma$.*

The condition means that f/dz is invariant, in which case f is said to be an inverse differential.

The definition (7.1) of f shows that $f(0) = 0$. If the condition in the lemma is fulfilled it follows that $f(A0) = 0$ for all $A \in \Gamma$. But the points $A0$ are dense on Σ and f is continuous. Therefore f vanishes on Σ, and Lemma 8 shows that $\nu \in N(\Gamma)$.

Consider now the function

$$P_A(z) = f(Az)A'(z)^{-1} - f(z). \tag{7.7}$$

By use of (7.3) we obtain

$$(P_A)_{\bar z} = \nu(Az)\bar{A}'(z)A'(z)^{-1} - \nu(z) = 0 \tag{7.8}$$

almost everywhere; in particular, also on Σ. Furthermore, the derivatives of P_A are locally square integrable (except perhaps at $A^{-1}\infty$). It follows that P_A is analytic, and if it vanishes on Σ it must be identically zero. This proves the lemma.

A more detailed investigation of the singularities at ∞ and $A^{-1}\infty$ shows that P_A is always a quadratic polynomial. It is of some interest to note the formula

$$P_{AB}(z) = P_A(Bz)B'(z)^{-1} + P_B(z). \tag{7.9}$$

7.4. The last lemma leads immediately to the principal result:

THEOREM 4. *If Γ is finitely generated, then $S = \Omega/\Gamma$ is of finite type.*

Indeed, all relations $f(Az) = f(z)A'(z)$ are consequences of the corresponding relations for the generators. On the other hand, each relation is equivalent to the vanishing of the coefficients of P_A, and thus to three linear conditions on ν. For a finitely generated group the space $B(\Gamma)/N(\Gamma)$ has thus finite dimension. By Lemma 3 and Theorem 3 this means that $Q^*(\Gamma)$ is finite dimensional, and hence the same is true of $Q(\Gamma)$. We apply Theorem 1 to conclude that S is of finite type.

8. Invariant regions.

8.1. It is possible to obtain more precise information by adding the hypothesis that one or several Ω_i are connected, and hence invariant under the full group Γ. Such groups have been called function groups.[4]

The fact that all invariant regions Ω_i have the same boundary Σ does not by itself preclude the existence of any number of such regions. However, the existence of nonelliptic fixpoints does impose a severe restriction. The following theorem was communicated to me by R. Accola.

THEOREM 5. *There are at most two invariant regions Ω_i, and if there are two they are simply connected.*

For the simple proof, which is purely topological, we refer to a forthcoming article by Accola.

8.2. If Γ is finitely generated we are able to prove a more precise result:

---

[4] L. Ford, Automorphic Functions (2nd ed., Chelsea, 1951), p. 64.

THEOREM 6. *Suppose that* Ω_1 *and* Ω_2 *are invariant regions for the finitely generated group* Γ. *Then* $S = \Omega/\Gamma$ *has only the components* $S_1 = \Omega_1/\Gamma$ *and* $S_2 = \Omega_2/\Gamma$. *Moreover,* S_1 *and* S_2 *are homeomorphic with isomorphic signatures.*

Since Γ is finitely generated we need not distinguish between $Q(\Gamma)$ and $Q^*(\Gamma)$. We shall denote by Q_i the subspace of quadratic differentials that vanish outside of Ω_i. The main part of the theorem is an easy consequence of

LEMMA 10. *If* Ω_1 *is connected, then* $\dim Q \leqq 2 \dim Q_1$.

Indeed, under the hypothesis of the theorem the lemma yields $\dim Q_1 + \dim Q_2 \leqq \dim Q \leqq 2 \dim Q_1$. Hence $\dim Q_2 \leqq \dim Q_1$, and by symmetry $\dim Q_1 = \dim Q_2$. Furthermore, $\dim Q = \dim Q_1 + \dim Q_2$, and this shows that there is no room for a third component. If we can show that S_1 and S_2 have the same signature it will follow, by the formula for the number of linearly independent quadratic differentials, that they have the same genus.

The points with finite signature arise from elliptic fixpoints, and it is easy to see that each elliptic transformation must have one fixpoint in Ω_1 and one in Ω_2 (the simplest way to see this is by conformal mapping of Ω_1 and Ω_2 on the unit disk). It is therefore evident that the points with finite signature can be matched in pairs, in such a way that corresponding points have the same signature.

The facts for points with infinite signature are quite similar. We know by Lemma 1 that any such point, whether on S_1 or S_2, determines a class of conjugate parabolic transformations. In the present situation the converse is also true. Let $A \in \Gamma$ be parabolic. We map Ω_1 conformally on the unit disk D and use the construction in 6.1 to determine an isomorphic mapping of Γ on a Fuchsian group G. The image B of A is necessarily parabolic. In the elementary case of a Fuchsian group it is quite evident that B corresponds to a point on S_1 with infinite signature (M. Heins has proved this for arbitrary Fuchsian groups).[5] Moreover, non-conjugate transformations correspond to different points. It follows that the points with infinite signature are indeed in one-one correspondence.

8.3. It remains to prove Lemma 10. Let us denote by Q'_1 the space of quadratic differentials which vanish on Ω_1. Given $\phi \in Q'_1$ we form $\nu = \bar{\phi}\rho^2$, which is a Beltrami differential, and construct

---

[5] M. Heins, "On Fuchsoid groups that contain parabolic transformations" (*Contributions to function theory*), Tata Institute, Bombay, 1960.

$$S\phi(\zeta) = \int_{\Omega-\Omega_1} \nu(z)\,(z-\zeta)^{-4}\,dxdy \tag{8.1}$$

for $\zeta \in \Omega_1$. One verifies that $S\phi$ is a quadratic differential. Because ν is bounded the integral is majorized by

$$\int_{|z-\zeta|>\delta(\zeta)} |z-\zeta|^{-4}\,dxdy$$

which is a multiple of $\delta(\zeta)^{-2}$. Since Ω_1 is simply connected $\delta(\zeta)^{-1}$ is comparable with $\rho(\zeta)$ (we are assuming that $\infty \in \Sigma$). It follows that $S\phi \in Q_1$.

We shall show that the antilinear mapping $S\colon Q'_1 \to Q_1$ is one to one. Assuming that $0, 1, \infty \in \Sigma$ we construct, as in 7.1,

$$f(\zeta) = -\pi^{-1}\int_{\Omega-\Omega_1} \nu(z)\left[\frac{1}{z-\zeta} - \frac{\zeta}{z-1} - \frac{1-\zeta}{z}\right]dxdy. \tag{8.2}$$

If $S\phi = 0$ we see by comparison with (8.1) that $f'''(\zeta) = 0$ in Ω_1. Hence $f(\zeta)$ is a quadratic polynomial in Ω_1. Its behavior at ∞ shows that it is in fact a linear polynomial, and because it vanishes at 0 and 1 it reduces identically to zero. By continuity $f = 0$ on Σ, and we conclude by Lemma 8 that $\nu \in N(\Gamma)$. Since $\nu = \bar{\phi}\rho^2$ this gives $\phi = 0$, and we have proved that S is one to one.

We emphasize that the connectedness of Ω_1 is essential for the proof. If Ω_1 were not connected we could merely conclude that f is a quadratic polynomial in each component, but not necessarily the same in all components.

HARVARD UNIVERSITY.

ON GENERALIZED ABELIAN INTEGRALS OF THE SECOND KIND AND MODULAR FORMS OF DIMENSION ZERO.

By MARVIN ISADORE KNOPP.

I. Introduction. This paper generalizes the results of [1] in a manner described below. Let $\Gamma(j)$, $j=2,3,\cdots$ denote the principal congruence subgroup of the modular group $\Gamma(1)$. By a *generalized abelian integral* (or GAI) for $\Gamma(j)$ we will mean here a function $f(\tau)$, regular in the upper half-plane, $\operatorname{Im}(\tau)>0$, and such that for $\operatorname{Im}(\tau)>0$,

$$f(V\tau)=v(V)f(\tau)+C(V), \text{ for } V\in\Gamma(j),$$

where $v(V)$ and $C(V)$ are independent of τ, and $|v(V)|=1$. The numbers $C(V)$ are the *periods* of f while v is a character on $\Gamma(j)$. We note that $f(\tau)$ is a GAI for $\Gamma(j)$ if and only if $\frac{d}{d\tau}f(\tau)$ is a modular form of dimension -2.

While in [1] the discussion was limited to the case $v\equiv 1$, that is, to the case of abelian integrals, we here carry over all of the results of [1] to the class of GAI's, subject to the rather natural restriction on the character v that there exists a positive integer n such that

$$(1.1) \qquad v(V)=1 \text{ for all } V\in\Gamma(nj).$$

Such a v is called a *congruence character.* This restriction on v is imposed to insure that certain exponential sums, which are closely related to Kloosterman sums, have a nontrivial asymptotic estimate (cf. § II. 3, esp. (2.6)).

If f has at most a polar singularity in the local uniformizing variable at each parabolic cusp (cf. § II. 1) of $\mathcal{F}_j$, a fundamental domain for $\Gamma(j)$, then we say that f is of the *second kind*; if f is regular in the local variable at each cusp, we say that f is of the *first kind.* If f is of the second kind and $C(V)=0$ for all $V\in\Gamma(j)$, then f is a modular form of dimension zero for $\Gamma(j)$. For convenience we introduce some notation. We will let $[\Gamma(j),v]_2$ be the vector space (over the complex numbers) of GAI's of the second kind; $[\Gamma(j),v]_1$ will denote those of the first kind; as usual $\{\Gamma(j),0,v\}$ will denote the space of modular forms of dimension zero for $\Gamma(j)$.

Received April 8, 1963.

Research supported in part by National Science Foundation grant number G-14362.

Our main results are the following.

(i) Construction of an explicit basis for $[\Gamma(j), v]_1$ (Theorem (4.1)).

(ii) Construction of an explicit basis for $[\Gamma(j), v]_2$ modulo $[\Gamma(j), v]_1$. This combined with (i) gives a basis for $[\Gamma(j), v]_2$ (Theorem (4.4)).

(iii) Explicit calculation of the Fourier coefficients of elements of $\{\Gamma(j), 0, v\}$ (Theorem (5.3)). These are of the same general form as the Fourier coefficients of automorphic forms of *positive* dimension on general H-groups which have been computed by the circle method (cf. [3]).

It is to be kept in mind that these results are obtained only for congruence characters v.

A key step in the derivation of (i) and (iii) is the use of a "gap theorem" for automorphic forms due to Petersson [4, pp. 460-1]. This is a generalization of the Weierstrass gap theorem which occurs in the classical theory of compact Riemann surfaces (cf. [6, p. 272]). The latter is sufficient for the case $v \equiv 1$ and was indeed used for this case in [1].

For the most part the results will be given without proof, the proofs being similar to those of corresponding statements in [1].

These results can be carried over to arbitrary congruence subgroups Γ of $\Gamma(1)$, that is, to groups Γ such that there is a j with $\Gamma(j) \subseteq \Gamma \subseteq \Gamma(1)$. The changes that have to be made are ones of detail only.

II. Preliminary remarks and definition of the basic functions.

1. Let M_α $(\alpha = 1, 2, \cdots)$ be the elements of $\Gamma(1)$, with $M_1 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Here we will not distinguish between the linear fractional transformation $\tau' = \frac{a\tau + b}{c\tau + d}$ and the matrices $\pm \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Let $\mathfrak{F}_j$ be a fundamental domain for $\Gamma(j)$ with ∞ as a cusp. Write $P_\alpha = M_\alpha^{-1}(\infty)$; in this way ge get all parabolic cusps P_α of $\Gamma(j)$ with $P_1 = \infty$. It is known that the subgroup of $\Gamma(j)$ leaving P_α fixed is a cyclic group with generator $T_\alpha = M_\alpha^{-1} S M_\alpha$, where $S = \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix}$. The set of P_α consists of all rational points on the real line and the point ∞. At a parabolic cusp $P = P_\alpha$, the *local uniformizing variable* is defined by

$$\begin{aligned} t_P = t_\alpha &= \exp\left[-\frac{2\pi i}{s^2 j(\tau - r/s)}\right], \text{ if } P = r/s < \infty, \quad (r, s) = 1 \\ &= \exp(2\pi i \tau / j), \text{ if } P = \infty. \end{aligned}$$

Let v be a character on $\Gamma(j)$, that is, a homomorphism of $\Gamma(j)$ into the unit circle of the complex plane. Note that $T_\alpha \in \Gamma(j)$ and write $v(T_\alpha) = \exp(2\pi i\kappa_\alpha)$, $0 \leqq \kappa_\alpha < 1$; for convenience write $v(T_1) = v(S) = \exp(2\pi i\kappa)$, $0 \leqq \kappa < 1$.

2. Let ν be a positive integer and let $M_\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a fixed element of $\Gamma(1)$. Then we define $\delta(M_\alpha)$ by

$$(2.1) \qquad \begin{aligned} \delta(M_\alpha) &= 1, \text{ if } c \equiv 0,\ a \equiv d \equiv \pm 1 \pmod{j} \\ &= 0, \text{ otherwise.} \end{aligned}$$

Put

$$(2.2) \qquad \begin{aligned} a_m(j, \nu, M_\alpha, v) = a_m(\nu, M_\alpha, v) &= \left(\frac{2\pi}{j}\right) \sum_{k=1}^{\infty} {}^{\alpha} k^{-1} A(m, k, \nu, M_\alpha, v) \\ &\cdot \left(\frac{\nu - \kappa_\alpha}{m + \kappa}\right)^{1/2} I_1\left(\frac{4\pi}{k} \frac{\sqrt{(m+\kappa)(\nu - \kappa_\alpha)}}{j}\right), \text{ if } m + \kappa > 0 \\ &= \frac{4\pi^2(\nu - \kappa_\alpha)}{j^2} \sum_{k=1}^{\infty} {}^{\alpha} k^{-2} A(0, k, \nu, M_\alpha, v), \text{ if } m = \kappa = 0, \end{aligned}$$

where

$$(2.3) \qquad \begin{aligned} A(m, k, \nu, M_\alpha, v) = &\sum_{\substack{0 \leqq h < jk \\ (h,k)=1}} {}^{\alpha}\, \bar{v}(M_\alpha{}^{-1} V_{k,-h}) \\ &\cdot \exp\left[-\frac{2\pi i}{jk}\{(\nu - \kappa_\alpha)h' + (m + \kappa)h\}\right] \end{aligned}$$

and I_1 is the modified Bessel function of the first kind. In (2.3) h' is any integral solution of $hh' \equiv -1 \pmod{k}$; we put $k' = -k^{-1}(hh' + 1)$ and define $V_{k,-h} = \begin{pmatrix} h' & k' \\ k & -h \end{pmatrix}$. The superscript α appearing in (2.2) and (2.3) indicates that the sums are restricted by the condition

$$(2.4) \qquad V_{k,-h} \equiv \pm M_\alpha \pmod{j},$$

where the congruence is elementwise. This condition is to be understood in the sense that each appropriate lower row $(k, -h)$ occurs *exactly once* and that we are then to find any (h', k') such that $-hh' - kk' = 1$ and such that (2.4) is satisfied. Of course, (2.4) implies that $M_\alpha{}^{-1} V_{k,-h} \in \Gamma(j)$, so that $v(M_\alpha{}^{-1} V_{k,-h})$ is defined. (The bar here denotes complex conjugation, as usual.)

The ambiguity in the choice of (h', k') leads to a possible ambiguity in the definition (2.3) since the summand in the right hand side of (2.3) depends upon h'. However (h', k') can be replaced only by $(h' + tjk, k' - tjh)$, with t an integer. Then $V_{k,-h}$ is replaced by

$$\begin{pmatrix} h' + tjk & k' - tjh \\ k & -h \end{pmatrix} = S^t V_{k,-h},$$

and

$$\bar{v}(M_\alpha^{-1} V_{k,-h}) \exp[-\frac{2\pi i}{jk}(\nu - \kappa_\alpha) h']$$

is replaced by

$$\bar{v}(M_\alpha^{-1} S^t V_{k,-h}) \exp[-\frac{2\pi i}{jk}(\nu - \kappa_\alpha)(h' + tjk)]$$
$$= \bar{v}(T_\alpha{}^t M_\alpha^{-1} V_{k,-h}) e^{2\pi i \kappa_\alpha t} \exp[-\frac{2\pi i}{jk}(\nu - \kappa_\alpha) h'],$$

so that the summand in (2.3) is well-defined.

3. We now define the basic functions

$$\lambda(j, \nu, M_\alpha, v; \tau) = \lambda(\nu, M_\alpha, v; \tau)$$

by

$$\lambda(\nu, M_\alpha, v; \tau) = \delta(M_\alpha) \bar{v}(M_\alpha^{-1} V_{c,d}) \exp[\frac{-2\pi i(\nu - \kappa_\alpha)}{j} V_{c,d}(\tau)]$$
$$\text{(2.5)} \qquad + \sum_{m=0}^{\infty} a_m(\nu, M_\alpha, v) \exp\{2\pi i(m + \kappa)\tau/j\},$$

for $\operatorname{Im}(\tau) > 0$, where $V_{c,d} \in \Gamma(1)$ has the same second row as M_α. The calculation used in the preceding paragraph to show that $A(m, k, \nu, M_\alpha, v)$ is well-defined proves that the first term on the right hand side of (2.5) is independent of the particular choice of $V_{c,d}$ and therefore is well-defined.

The functions λ are regular in the upper half-plane, $\operatorname{Im}(\tau) > 0$. The proof of this fact uses the nontrivial asymptotic estimate

$$\text{(2.6)} \qquad A(m, k, \nu, M_\alpha, v) = O(k^{\frac{1}{2}+\epsilon}), \ \epsilon > 0, \text{ as } k \to +\infty.$$

The constant involved in the estimate can be chosen independent of m. That (2.6) holds for all sums $A(m, k, \nu, M_\alpha, v)$, with v a congruence character (that is, satisfying (1.1)), was shown by Petersson [5, pp. 16-19]. It is worth noting that any estimate of the type

$$\text{(2.7)} \qquad A(m, k, \nu, M_\alpha, v) = O(k^{1-\delta}), \ \delta > 0, \text{ as } k \to +\infty$$

would suffice to carry through our methods, and in fact, all of the results presented hold if v satisfies (2.7). The estimate (2.6) is critical in the proof of the following fundamental result (cf. [2], for details).

THEOREM (2.8). (i) *Let* $V=\begin{pmatrix} * & * \\ \gamma & * \end{pmatrix}\in\Gamma(1)$. *Then, providing* v *satisfies* (1.1), *we have, for* $\operatorname{Im}(\tau)>0$,

$$\lambda(\nu, M_\alpha, v; V\tau)=\lambda(\nu, M_\alpha V, v^*;\tau)+\omega(M_\alpha, V), \tag{2.9}$$

where $\omega(M_\alpha, V)$ *is independent of* τ, *and* v^* *is the character on* $\Gamma(j)$ *defined by* $v^*(T)=v(VTV^{-1})$, *for* $T\in\Gamma(j)$. (*Note that* v^* *satisfies* (1.1), *if* v *does.*)

(ii) *If* $\gamma=0$, *then* $\omega(M_\alpha, V)=0$, *and if* $\gamma\neq 0$, *then*

$$\omega(M_\alpha, V)=\rho(\nu, M_\alpha, v, V)+\sigma(\nu, M_\alpha, v, V);$$

here, letting $M_\alpha V=M_\beta$,

$$\begin{aligned}\sigma(\nu, M_\alpha, v, V) \\ =\lim_{K\to\infty}\Bigg\{ & \sum_{\substack{0<k\leq[tK]\\(k,m)=1}}\ \sum_{|m|\leq K}{}^{\beta}\ \bar{v}^*(M_\beta{}^{-1}V_{k,-m})\exp\left\{\frac{-2\pi i(\nu-\kappa_\alpha)m'}{jk}\right\} \\ & -\sum_{\substack{(k,m)\in\mathcal{T}(K)\\(k,m)=1}}\sum{}^{\alpha}\ \bar{v}(M_\alpha{}^{-1}V_{k,-m})\exp\{-2\pi i(\nu-\kappa_\alpha)m'/jk\}\Bigg\},\end{aligned} \tag{2.10}$$

where t *is a certain positive constant and* $\mathcal{T}(K)$ *is a certain trapezoid in the* $k-m$ *plane* (t *and* $\mathcal{T}(K)$ *depend upon* V), *and*

$$\rho(\nu, M_\alpha, v, V)=\varphi(\nu, M_\alpha, v)-\varphi(\nu, M_\beta, v^*). \tag{2.11}$$

In (2.11), *we have put*

$$\begin{aligned}\varphi(\nu, M_\alpha, v)&=\tfrac{1}{2}a_0(\nu, M_\alpha, v),\ \text{if } \kappa=0\\ &=0 \qquad\qquad\quad ,\ \text{if } \kappa>0.\end{aligned} \tag{2.12}$$

III. Generalized abelian integrals. 1. In Theorem (2.8), if we let $V\in\Gamma(j)$, then we observe from (2.2)-(2.5) that $v^*=v$, and $\lambda(\nu, M_\alpha, v^*;\tau)=v(V)\lambda(\nu, M_\alpha, v;\tau)$. Hence applying the reasoning of the proof of Theorem (3.10) in [1], we obtain

THEOREM (3.1). *If* v *satisfies* (1.1), *then for*

$$M_\alpha=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in\Gamma(1),\ \lambda(\nu, M_\alpha, v;\tau)$$

is in $[\Gamma(j), v]_2$ *and has precisely one pole of order* $\nu-\kappa_\alpha$ *in* $\mathfrak{F}_j$. *That is,* λ *is regular in* $\operatorname{Im}(\tau)>0$ *and*

(i) *for* $V=\begin{pmatrix} * & * \\ \gamma & * \end{pmatrix}\in\Gamma(j)$, $\operatorname{Im}(\tau)>0$, *if* $\gamma\neq 0$

$$\begin{aligned}\lambda(\nu, M_\alpha, v; V\tau)-v(V)\lambda(\nu, M_\alpha, v;\tau) &\\ &=\sigma(\nu, M_\alpha, v, V)+\rho(\nu, M_\alpha, v, V)\end{aligned}$$

$$=\begin{cases}\sigma(\nu, M_\alpha, v, V), & \text{if } \kappa>0, \\ \sigma(\nu, M_\alpha, v, V)+\tfrac{1}{2}a_0(\nu, M_\alpha, v)(1-v(V)), & \text{if } \kappa=0,\end{cases} \tag{3.2}$$

with σ *defined by* (2.10) *and* ρ *by* (2.11)-(2.12), *and if* $\gamma=0$

$$\lambda(\nu, M_\alpha, v; V\tau)=v(V)\lambda(\nu, M_\alpha, v;\tau),$$

(ii) *with* $\mathfrak{F}_j$ *chosen such that* P_α *is a parabolic cusp of* $\mathfrak{F}_j$, *about this cusp* λ *has an expansion of the form*

$$e(\nu,\alpha)t_\alpha^{-(\nu-\kappa_\alpha)}+\sum_{m=0}^{\infty}b_m t_\alpha^{m+\kappa_\alpha}, \tag{3.3}$$

where

$$\begin{aligned} e(\nu,\alpha)&=\exp(-2\pi i(\nu-\kappa_\alpha)b/j), \text{ if } c=0 \\ &=\exp(-2\pi i(\nu-\kappa_\alpha)a/jc), \text{ if } c\neq 0,\end{aligned} \tag{3.4}$$

(iii) *at any other cusp* P_ι *of* $\mathfrak{F}_j$, λ *has an expansion of the form*

$$\sum_{m=0}^{\infty}b_m t_\iota^{m+\kappa_\iota}. \tag{3.5}$$

2. We now introduce the *series supplementary to* $\lambda(\nu, M_\alpha, v;\tau)$. Recall that ν is a *positive* integer and define ν' by

$$\begin{aligned}\nu'&=1-\nu, \text{ if } \kappa_\alpha>0 \\ &=-\nu, \quad \text{if } \kappa_\alpha=0.\end{aligned}$$

Let $a_m(\nu', M_\alpha, \bar{v})$ be the result of replacing v by $\bar{v}$ and ν by ν' in (2.2). Note that κ_α is then replaced by $\kappa'_\alpha=1-\kappa_\alpha$, if $\kappa_\alpha>0$ and by $\kappa'_\alpha=0$ if $\kappa_\alpha=0$. The same is true of κ' which replaces κ. We define $\hat{\lambda}$ by

$$\begin{aligned}\hat{\lambda}(\nu', M_\alpha, \bar{v};\tau)&=\delta(M_\alpha)v(M_\alpha^{-1}V_{c,d})\exp\Big[\frac{-2\pi i(\nu'-\kappa'_\alpha)}{j}V_{c,d}(\tau)\Big] \\ &\quad+\sum_{m+\kappa'>0}a_m(\nu', M, \bar{v})\exp\{2\pi i(m+\kappa')\tau/j\}.\end{aligned} \tag{3.6}$$

Notice that in contrast to the definition (2.5) of λ, the series (3.6) has no term corresponding to $m=0$, if $\kappa=0$. The definition is given this way as a matter of convenience and has the effect of simplifying some of the succeeding calculations.

We observe that the proof of Theorem (2.8) does not in any way depend upon the fact that ν is positive. Hence Theorem (2.8) applies to $\hat{\lambda}$ and we can deduce the counterpart of Theorem (3.1) for $\hat{\lambda}$. It is, namely

THEOREM (3.7). *If $\bar{v}$ satisfies* (1.1), *then for $M_\alpha \in \Gamma(1)$, $\lambda(\nu', M_\alpha, \bar{v}; \tau)$ is in $[\Gamma(j), v]_1$. That is, $\hat{\lambda}$ is regular in* $\operatorname{Im}(\tau) > 0$, *satisfies the transformation equations* (3.2) *of Theorem* (3.1), *with $\sigma(\nu, M_\alpha, v, V)$, and $\rho(\nu, M_\alpha, v, V)$ replaced by $\sigma(\nu', M_\alpha, \bar{v}, V)$ and $-\rho(\nu', M_\alpha, \bar{v}, V)$, respectively, and at each parabolic cusp of $\mathcal{F}_j$, $\hat{\lambda}$ has an expansion of the form* (3.5).

We see that $\nu' - \kappa'_\alpha = -(\nu - \kappa_\alpha)$ and therefore by (2.10), (2.11)-(2.12), and (2.2)-(2.3)

$$\sigma(\nu', M_\alpha, \bar{v}, V) = \bar{\sigma}(\nu, M_\alpha, v, V) \tag{3.8}$$

$$\begin{aligned} \rho(\nu', M_\alpha, \bar{v}, V) &= 0, \text{ if } \kappa > 0 \\ &= -\bar{\rho}(\nu, M_\alpha, v, V), \text{ if } \kappa = 0. \end{aligned} \tag{3.9}$$

If now $\Lambda(\tau)$ is a finite linear combination of the functions λ, say

$$\Lambda(\tau) = \sum_\alpha \sum_\nu b(\nu, \alpha) \lambda(\nu, M_\alpha, v; \tau),$$

we define $\hat{\Lambda}(\tau)$, the *function supplementary* to $\Lambda(\tau)$, by

$$\hat{\Lambda}(\tau) = \sum_\alpha \sum_\nu \bar{b}(\nu, \alpha) \hat{\lambda}(\nu', M_\alpha, \bar{v}; \tau),$$

We have the following fundamental relationship between $\Lambda(\tau)$ and $\hat{\Lambda}(\tau)$.

THEOREM (3.10). $\Lambda(\tau) \in \{\Gamma(j), 0, v\}$ *if and only if* $\hat{\Lambda}(\tau) \equiv 0$.

Proof. For $\kappa > 0$, by Theorem (3.1), if $V = \begin{pmatrix} * & * \\ \gamma & * \end{pmatrix} \in \Gamma(j)$,

$$\begin{aligned} \Lambda(V\tau) - v(V)\Lambda(\tau) &= \sum_\alpha \sum_\nu b(\nu, \alpha) \sigma(\nu, M_\alpha, v, V), && \text{if } \gamma \neq 0 \\ &= 0, && \text{if } \gamma = 0. \end{aligned}$$

On the other hand, by Theorem (3.7) and (3.8)-(3.9),

$$\begin{aligned} \hat{\Lambda}(V\tau) - \bar{v}(V)\hat{\Lambda}(\tau) &= \sum_\alpha \sum_\nu \bar{b}(\nu, \alpha) \bar{\sigma}(\nu, M_\alpha, v, V), && \text{if } \gamma \neq 0, \\ &= 0, && \text{if } \gamma = 0. \end{aligned}$$

Hence $\Lambda(\tau) \in \{\Gamma(j), 0, v\}$ if and only if $\hat{\Lambda}(\tau) \in \{\Gamma(j), 0, \bar{v}\}$. But $\hat{\Lambda}(\tau)$ is bounded in $\mathfrak{F}_j$ and hence is a modular form of dimension zero only if it is constant. By (3.6) the constant must be zero.

If $\kappa = 0$, then by Theorem (3.1),

$$\begin{aligned} \Lambda(V\tau) - v(V)\Lambda(\tau) &= \sum_{\alpha}\sum_{\nu} \mathrm{b}(\nu, \alpha)\sigma(\nu, M_\alpha, v, V) \\ &\quad + \tfrac{1}{2}(1 - v(V)) \sum_{\alpha}\sum_{\nu} b(\nu, \alpha) a_0(\nu, M_\alpha, v), \text{ if } \gamma \neq 0 \\ &= 0, \text{ if } \gamma = 0. \end{aligned}$$

By Theorem (3.7) and (3.8)-(3.9),

$$\begin{aligned} \hat{\Lambda}(V\tau) - v(V)\hat{\Lambda}(\tau) &= \sum_{\alpha}\sum_{\nu} \bar{b}(\nu, \alpha)\bar{\sigma}(\nu, M_\alpha, v, V) \\ &\quad + \tfrac{1}{2}(1 - \bar{v}(V)) \sum_{\alpha}\sum_{\nu} \bar{b}(\nu, \alpha) \bar{a}_0(\nu, M_\alpha, v), \text{ if } \gamma \neq 0 \\ &= 0, \text{ if } \gamma = 0. \end{aligned}$$

Hence, again $\Lambda(\tau) \in \{\Gamma(j), 0, v\}$ if and only if $\hat{\Lambda}(\tau) \in \{\Gamma(j), 0, \bar{v}\}$ and this happens if and only if $\hat{\Lambda}(\tau) \equiv 0$, by (3.6).

IV. Basis theorems. 1. In order to obtain our principal results we need the following important generalization, due to Petersson [4, pp. 460-1], of the classical Weierstrass gap theorem.

PETERSSON GAP THEOREM. *Let μ be the linear dimension over the complex numbers of the space of cusp forms in $\{\Gamma(j), -2, v\}$. Then given any parabolic cusp P_α of $\mathfrak{F}_j$, there exist exactly μ integers w_i, $0 < w_1 < \cdots < w_\mu$, depending on P_α, such that there does not exist an element of $\{\Gamma(j), 0, \bar{v}\}$, having as its only singularity in $\mathfrak{F}_j$ a pole of order $w_i - \kappa'_\alpha$ at P_α, $1 \leqq i \leqq \mu$.*

Remark. This is actually a very special case of Petersson's theorem, which deals with a very wide class of groups and automorphic forms of any *complex* dimension. Furthermore the restriction to parabolic points is unnecessary and has been made here simply to avoid the problem of defining the order of a pole at an arbitrary point of $\mathfrak{F}_j$. Note that we have defined the order of a pole at a parabolic cusp P_α in terms of the expansion of the function at that cusp. It is not an integer unless $\kappa'_\alpha = 0$.

We fix our attention on a single, arbitrarily chosen parabolic cusp P_h of a fixed fundamental domain $\mathfrak{F}_j$ with ∞ as a cusp. Write $P_h = r/s$, $(r, s) = 1$ (possibly $P_h = 1/0 = \infty$). Then there are integers t, u such that $M=\begin{pmatrix} t & u \\ s & -r \end{pmatrix}\in\Gamma(1)$, and then $M_k^{-1}(\infty)=P_k$. By Theorem (3.1), since $\bar{v}$ satisfies (1.1) if and only if v does, $\lambda(\nu, M_k, \bar{v}; \tau)$ is in $[\Gamma(j), \bar{v}]_2$, and has a pole of order $\nu-\kappa'_k$ at P_k and no other singularity in $\mathfrak{F}_j$. Consider the functions $\lambda(w_i, M_k, \bar{v}; \tau)=\lambda_i(\tau)$, where w_i $(1\leqq i\leqq\mu)$ is the sequence of integers connected with the cusp P_k in the Petersson gap theorem.

THEOREM (4.1). *The functions* $\hat{\lambda}(w'_i, M_k, v; \tau)=\hat{\lambda}_i(\tau)$ $(1\leqq i\leqq\mu)$ *are linearly independent over the complex numbers and, together with the constant function* 1, *they form a basis for* $[\Gamma(j), v]_1$.

Proof. By Theorem (3.7) each $\hat{\lambda}_i\in[\Gamma(j), v]_1$. If

$$\hat{g}(\tau)=\sum_{i=1}^{\mu} b_i\hat{\lambda}_i(\tau)+b_{\mu+1}\equiv 0,$$

then by (3.6), $b_{\mu+1}=0$, and by Theorem (3.10),

$$g(\tau)=\sum_{i=1}^{\mu}\bar{b}_i\lambda_i(\tau)\in\{\Gamma, 0, \bar{v}\}.$$

Unless $b_i=0$ for all $i=1,\cdots,\mu$, $g(\tau)$ is a modular form of the type excluded by the Petersson gap theorem. Hence $b_1=\cdots=b_\mu=b_{\mu+1}=0$.

But since the space of cusp forms in $\{\Gamma(j), -2, v\}$ has dimension μ, $[\Gamma, v]_1$ has dimension $\mu+1$ and therefore $\hat{\lambda}_1,\cdots,\hat{\lambda}_\mu, 1$ form a basis for $[\Gamma(j), v]_1$.

2. Let $P_1=\infty, P_2,\cdots, P_q$ be the parabolic cusps of $\mathfrak{F}_j$. It is known that $q=[\Gamma(1):\Gamma(j)]/j$. Suppose $f(\tau)\in[\Gamma(j), v]_2$. Then a linear combination of the functions $\lambda(\nu, M_\alpha, v; \tau)$ $(\alpha=1,\cdots,q; \nu>0)$ can be so constructed that its principal part agrees with the principal of $f(\tau)$ at each parabolic cusp. That is, there exist constants $b(\nu,\alpha)$ such that

$$f(\tau)-\sum_{\alpha=1}^{q}\sum_{\nu} b(\nu,\alpha)\lambda(\nu, M_\alpha, v; \tau)\in\Gamma(j), v]_1. \tag{4.2}$$

By Theorem (4.1) there are constants b_i $(1\leqq i\leqq\mu+1)$ such that

$$\begin{aligned} f(\tau)=\sum_{\alpha=1}^{q}\sum_{\nu} b(\nu,\alpha)\lambda(\nu, M_\alpha, v; \tau)& \\ +\sum_{i=1}^{\mu} b_i\hat{\lambda}(w'_i, M_k, v; \tau)+b_{\mu+1}.& \end{aligned} \tag{4.3}$$

That is, we have

THEOREM (4.4). *The functions* $\lambda(j, v, M_\alpha, v:\tau)$ $(1\leqq\alpha\leqq q, \nu>0)$ *to-*

gether with $1, \hat{\lambda}(w'_1, M_h, v; \tau), \cdots, \hat{\lambda}(w'_\mu, M_h, v; \tau)$ *form a basis for* $[\Gamma(j), v]_2$. *Here* P_h *is any parabolic cusp of* $\mathfrak{F}_j$ *and* $w_1, \cdots, w_\mu$ *is the gap sequence at* P_h.

V. Modular forms. 1. An element of $\{\Gamma(j), 0, v\}$ is in particular, in $[\Gamma(j), v]_2$ and can therefore be expressed in the form (4.3). We conclude that in this case $b_1, \cdots, b_\mu$ are all zero. That is, we have

THEOREM (5.1). *If* $f(\tau) \in \{\Gamma(j), 0, v\}$, *then*

$$f(\tau) = \sum_{\alpha=1}^{q} \sum_{\nu} b(\nu, \alpha) \lambda(\nu, M_\alpha, v; \tau) + b_{\mu+1}, \tag{5.2}$$

where the linear combination on the right hand side of (5.2) *is so constructed that its principal part agrees with that of* $f(\tau)$ *at each parabolic cusp of* $\mathfrak{F}_j$. *Furthermore, if* v *is not identically* 1 *on* $\Gamma(j)$, $b_{\mu+1} = 0$.

Proof. With $f(\tau)$ expressed as in (4.3), let

$$\hat{f}(\tau) = \sum_{\alpha=1}^{q} \sum_{\nu} \bar{b}(\nu, \alpha) \hat{\lambda}(\nu', M_\alpha, \bar{v}; \tau) + \sum_{i=1}^{\mu} \bar{b}_i \lambda(w_i, M_h, \bar{v}; \tau) + \bar{b}_{\mu+1}.$$

The calculations used in the proof of Theorem (3.10) show that $f(\tau) \in \{\Gamma(j), 0, v\}$ if and only if $\hat{f}(\tau) \in \{\Gamma(j), 0, \bar{v}\}$. If some $\bar{b}_i \neq 0$ $(1 \leqq i \leqq \mu)$, $\hat{f}(\tau)$ is a modular form of the type excluded by the Petersson gap theorem. Therefore, $b_1 = b_2 = \cdots = b_\mu = 0$.

Furthermore, $\hat{f}(\tau) \in \{\Gamma(j), 0, \bar{v}\}$ is then bounded in the upper half-plane and therefore a constant. By the definition (3.6) of $\hat{\lambda}$, we conclude that $\hat{f}(\tau) = \bar{b}_{\mu+1}$. If $v \not\equiv 1$ on $\Gamma(j)$, $b_{\mu+1} = 0$, since a nonzero constant is not in $\{\Gamma(j), 0, \bar{v}\}$ in this case.

2. We are now in a position to read off the Fourier coefficients of $f(\tau) \in \{\Gamma(j), 0, v\}$ in its expansion about ∞, that is, in powers of $\exp(2\pi i\tau/j)$ directly from (2.2)-(2.3) and Theorem (5.1). The result is

THEOREM (5.3). *Let* $f(\tau) \in \{\Gamma(j), 0, v\}$ *and let the principal part of* $f(\tau)$ *at* P_α *be* $\sum_{\nu} b(\nu, \alpha) t_\alpha^{-(\nu-\kappa_\alpha)}$, *for* $1 \leqq \alpha \leqq q$. *Write the Fourier expansion of* $f(\tau)$ *in the variable* $t_1 = \exp(2\pi i\tau/j)$ *as*

$$f(\tau) = \sum_{\nu} b(\nu, 1) e^{-2\pi i(\nu-\kappa)\tau/j} + \sum_{m \geqq 0} a_m e^{2\pi i(m+\kappa)\tau/j}.$$

Then, if $v \not\equiv 1$ *on* $\Gamma(j)$, *the coefficients* a_m, $m \geqq 0$, *are given by*

$$a_m = \sum_{\alpha=1}^{q} \sum_{\nu} b(\nu, \alpha) \varepsilon(\nu, \alpha) a_m(\nu, M_\alpha, v), \tag{5.4}$$

where $a_m(\nu, M_\alpha, v)$ *is defined by* (2.2) *and* $e(\nu,\alpha)$ *by* (3.4). *If* $v \equiv 1$ *on* $\Gamma(j)$, *the* a_m *are determined in the same way for* $m \geqq 1$.

We could obtain the expansion about each of the other cusps of $\mathfrak{F}_j$ by applying Theorem (2.8), part (i) to $f(\tau)$.

THE UNIVERSITY OF WISCONSIN,
MADISON, WISCONSIN.

REFERENCES.

[1] M. I. Knopp, "On abelian integrals of the second kind and modular functions," *American Journal of Mathematics*, vol. 84 (1962), pp. 615-628.

[2] ———, "Construction of automorphic forms on H-groups and supplementary Fourier series," *Transactions of the American Mathematical Society*, vol. 103 (1962), pp. 168-188.

[3] J. Lehner, "The Fourier coefficients of automorphic forms on horocyclic groups, II," *Michigan Mathematical Journal*, vol. 6 (1959), pp. 173-193.

[4] H. Petersson, "Über eine Metrisierung der automorphen Formen und die Theorie der Poincaréschen Reihen," *Mathematische Annalen*, vol. 117 (1940), pp. 453-537.

[5] ———, "Über Modulfunktionen und Partitionenprobleme," *Abhandlungen Deutschen Akademie der Wissenschaften zu Berlin, Klasse für Mathematik und allgemeine Naturwissenschaften*, 1954, no. 2, 59 pp.

[6] G. Springer, *Introduction to Riemann surfaces*, Reading, Addison-Wesley, 1957.

A CLASS OF SINGULAR INTEGRALS.*

By B. Frank Jones, Jr.

1. Introduction. A. P. Calderón and A. Zygmund [1] initiated a study of singular integrals which generalize to n-dimensional space many results concerning the classical Hilbert transform. Their results have had wide application to partial differential equations, partly due to the fact that certain derivatives of fundamenal solutions of Laplace's equation satisfy the conditions required by Calderón and Zygmund for kernels in the singular integral operators. The present article treats a different class of singular integrals; these generalize integral operators with kernels which arise from parabolic differential equations with constant coefficients. We now turn to an example of such a situation. The analogy with the case of Laplace's equation was described in [2].

Let $P(\xi)=P(\xi_1,\cdots,\xi_n)$ be a homogeneous polynomial of even degree m, such that $P(\xi)$ has negative real part for real ξ. Let $D=(\partial/\partial x_1,\cdots,\partial/\partial x_n)$. Then the parabolic equation

$$\partial u/\partial t=(-1)^{m/2}P(D)u$$

has for a fundamental solution the function

$$\Gamma(x,t)=(2\pi)^{-n}\int_{R^n}\exp[ix\cdot\xi+tP(\xi)]d\xi,\quad t>0$$

(cf. [4], pp. 97-102). Now let

$$k(x,t)=D_x{}^\rho\Gamma(x,t)=\frac{\partial^{\rho_1}}{\partial x_1}\cdots\frac{\partial^{\rho_n}}{\partial x_n}\Gamma(x,t),$$

where $\rho=(\rho_1,\cdots,\rho_n)$ is fixed and $\rho_1+\cdots+\rho_n=m$. Then

$$(1.1)\qquad k(x,t)=(2\pi)^{-n}\int_{R^n}(i\xi)^\rho\exp[ix\cdot\xi+tP(\xi)]d\xi,$$

Received October 17, 1963.

Revised January 31, 1964.

* This research was supported in part by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract No. AFOSR 62 233. Reproduction in whole or in part is permitted for any purpose of the United States Government.

where $(i\xi)^\rho = (i\xi_1)^{\rho_1} \cdots (i\xi_n)^{\rho_n}$. Now let us make the change of variable $y = t^{1/m}\xi$ in (1.1). Then

$$(1.2) \qquad k(x,t) = (2\pi)^{-n} t^{-1-n/m} \int_{E^n} (iy)^\rho \exp[it^{-1/m} x \cdot y + P(y)]\, dy.$$

We therefore obtain for $k(x,t)$ the relation

$$(1.3) \qquad k(\alpha^{1/m}x, \alpha t) = \frac{1}{\alpha^{1+n/m}} k(x,t), \quad \alpha > 0.$$

Note that this relation could also be expressed in the form

$$(1.4) \qquad k(x,t) = \frac{1}{t^{1+n/m}} k\left(\frac{x}{t^{1/m}}, 1\right), \quad t > 0,$$

which follows immediately from either (1.2) or (1.3).

For example, for the heat equation $\partial u/\partial t = \partial^2 u/\partial x^2$,

$$k(x,t) = \frac{1}{t^{3/2}} \frac{1}{\sqrt{\pi}} \left(\frac{x^2}{8t} - \frac{1}{4}\right) e^{-x^2/4t}.$$

From estimates on Γ ([4], Theorem 39) it follows that any derivative of Γ decays as fast as $\exp[-c\,|x|^{m/(m-1)}]$ as $|x| \to \infty$, where c is a positive constant; here $|x| = (x_1^2 + \cdots + x_n^2)^{1/2}$. Therefore, as k is a spatial derivative of an absolutely integrable function, it follows immediately that

$$(1.5) \qquad \int_{E^n} k(x,1)\, dx = 0.$$

We can express (1.3) by saying that k is homogeneous in a certain sense, and (1.5) by saying that k has mean value zero on hyperplanes orthogonal to the t-axis. As was pointed out in [2], these two properties are entirely analogous to corresponding homogeneity and mean value properties assumed in [1].

In dealing with parabolic equations it is frequently necessary to solve the equation

$$(1.6) \qquad \partial u/\partial t - (-1)^{m/2} P(D) u = f,$$

where $f(x,t)$ is a given function defined for (say) all x in E^n and all positive t. A particular solution of (1.6) may be expressed in the form

$$u(x,t) = \int_0^t \int_{E^n} \Gamma(x-\xi, t-\tau) f(\xi,\tau)\, d\xi d\tau,$$

in case f is (say) Hölder-continuous and has compact support. This solution satisfies the initial condition $u(x,0) \equiv 0$. Using the same notation as above, it then follows that ([3], Théorème 3)

$$(1.7)\qquad D_x{}^p u(x,t)=\lim_{\epsilon\to 0}\int_0^{t-\epsilon}\int_{R^n}k(x-\xi,t-\tau)f(\xi,\tau)\,d\xi d\tau.$$

Note that the function $k(x-\xi,t-\tau)f(\xi,\tau)$ actually fails to be integrable over the set $E=R^n\times(0,\infty)$. Hence, the integral in (1.7) must be taken in the indicated sense. In our case of Hölder-continuous f the limit in (1.7) exists and, moreover, the iterated integral

$$\int_0^t\Big[\int_{R^n}k(x-\xi,t-\tau)f(\xi,\tau)\,d\xi\Big]d\tau$$

exists in the usual sense.

The principal result of this paper implies that if $f\in L_p(E)$, where $1<p<\infty$, then the integral in (1.7) converges in $L_p(E)$ to a function f_0, and $\| f_0\|_{L_p(E)}\leqq A_p\| f\|_{L_p(E)}$, where the constant A_p depends only on k and p. This is a direct analog of a basic result in [1]. The proof is quite similar to the corresponding proof in [1], and we shall often refer to that paper in the development which follows.

Of course, an immediate application of the theorem is the *a priori* estimate $\| D_x{}^p u\|_{L_p(E)}\leqq A_p\| f\|_{L_p(E)}$ for the solution of (1.6) satisfying $u(x,0)\equiv 0$.

2. Description of the singular integrals. In this section a class of kernels will be described, generalizing those introduced in Section 1 connected with parabolic differential equations. We make the following assumptions:

I. $\phi(t)$ is a non-decreasing positive continuous function for $0<t<\infty$ such that $\lim_{t\to 0}\phi(t)=0$, $\lim_{t\to\infty}\phi(t)=\infty$;

II. $\Omega(x)$ is a measurable complex-valued function on R^n and

$$k(x,t)=\frac{1}{t\phi(t)^n}\Omega\left(\frac{x}{\phi(t)}\right),\quad t>0;$$

III. $\int_{R^n}(1+|x|)|\Omega(x)|\,dx<\infty$ and $\int_{R^n}\Omega(x)\,dx=0.$

These are certainly fundamental assumptions on the nature of $k(x,t)$; II and III correspond to the relations (1.4) and (1.5), respectively. The remaining assumptions are more or less technical in nature.

We require the following of $\phi(t)$. There exists a positive constant c such that:

IV. $\int_0^a\frac{\phi(t)}{t}\,dt\leqq c\phi(a),\quad 0<a<\infty,$

$$\int_a^\infty \frac{1}{t\phi(t)}\,dt \leqq \frac{c}{\phi(a)}, \quad 0 < a < \infty;$$

V. $\dfrac{\phi(2t)}{\phi(t)} \leqq c, \quad 0 < t < \infty;$

VI. for $0 < t < \infty$ and $|\delta| \leqq 1/2$,

$$\left|\frac{\phi((1+\delta)t)}{\phi(t)} - 1\right| \leqq c|\delta|.$$

Finally, we assume that $\Omega(x)$ satisfies the following for some positive c:

VII. $\displaystyle\int_{R^n} |\Omega(x) - \Omega(x+y)|\,dx \leqq c|y|, \quad y \in R^n;$

VIII. $\displaystyle\int_{R^n} |\Omega(x) - \Omega((1+\delta)x)|\,dx \leqq c|\delta|$ for $|\delta| \leqq 1/2;$

IX. $\displaystyle\int_{|x|\geqq a} |\Omega(x)|\,dx \leqq ca^{-n}$ for $a > 0$.

It should be noted that all of these assumptions are almost trivially fulfilled in the case of the kernel $k(x,t)$ treated in Section 1, where $\phi(t) = t^{1/m}$. Also, if $\phi(t) \equiv t^\beta$ for some positive β, then it is seen immediately that the conditions IV, V, VI are trivially satisfied.

Throughout this paper c shall stand for a (large) positive generic constant, which depends only on the kernel k and the dimension n of the space.

A formula for a change of dummy variable in a certain type of integral will frequently be needed; for convenience in later manipulations the formula will now be stated. If $0 < a < b \leqq \infty$ and $\mathcal{E} \subset R^n$, then

$$\int_a^b d\tau \int_{\xi\in\mathcal{E}} k(\xi,\tau)F(\xi,\tau)\,d\xi = \int_a^b \frac{d\tau}{\tau}\int_{\xi\in\frac{1}{\phi(\tau)}\mathcal{E}} \Omega(\xi)F(\phi(\tau)\xi,\tau)\,d\xi. \tag{2.1}$$

This relation follows immediately by using the expression II for $k(\xi,\tau)$ and replacing ξ on the left side of (2.1) by $\xi' = \dfrac{1}{\phi(\tau)}\xi$, and then replacing ξ' by ξ. Here $F(\xi,\tau)$ stands for various functions which will enter the discussion.

Now some truncations of k will be defined: if $\epsilon > 0$, let

$$k_\epsilon(x,t) = \begin{cases} k(x,t), & t \geqq \epsilon, \\ 0, & t < \epsilon. \end{cases} \tag{2.2}$$

If $0 < \epsilon < T < \infty$, let

$$k_{\epsilon,T}(x,t) = \begin{cases} k_\epsilon(x,t), & t \leqq T, \\ 0, & t > T. \end{cases} \tag{2.3}$$

We next define two convolution operators corresponding to these kernels. Recall that $E = R^n \times (0, \infty)$. For a function $f(x,t)$ defined on E let

$$f_\epsilon(x,t) = \int_0^\infty \int_{R^n} k_\epsilon(x-\xi, t-\tau) f(\xi,\tau)\, d\xi d\tau \tag{2.4}$$

and

$$f_{\epsilon,T}(x,t) = \int_0^\infty \int_{R^n} k_{\epsilon,T}(x-\xi, t-\tau) f(\xi,\tau)\, d\xi d\tau. \tag{2.5}$$

Note that $f_{\epsilon,T}(x,t) \equiv f_\epsilon(x,t)$ if $t \leqq T$. Also note that $f_\epsilon(x,t) = 0$ if $t \leqq \epsilon$. We now state a simple but convenient fact.

LEMMA 1. *The function $k_{\epsilon,T}$ is in $L_1(E)$.*

Proof. By (2.1) and III

$$\begin{aligned}\int_0^\infty \int_{R^n} |k_{\epsilon,T}(x,t)|\, dxdt &= \int_\epsilon^T \int_{R^n} |k(x,t)|\, dxdt \\ &= \int_\epsilon^T \frac{d\tau}{\tau} \int_{R^n} |\Omega(x)|\, dx \\ &\leqq c \log \frac{T}{\epsilon}.\end{aligned}$$

Q. E. D.

As a result of this lemma we have

THEOREM 1. *Let $1 < p < \infty$ and let $f \in L_p(E)$. Then $f_{\epsilon,T}(x,t)$ exists as an absolutely convergent integral almost everywhere in E and $f_{\epsilon,T} \in L_p(E)$. Also, $f_\epsilon(x,t)$ exists as an absolutely convergent integral almost everywhere in E and is locally in L_p.*

Proof. By W. H. Young's inequality ([6], p. 37) and Lemma 1 the assertions for $f_{\epsilon,T}$ follow immediately. The assertions for f_ϵ follow from the fact that $f_{\epsilon,T} \equiv f_\epsilon$ for $t \leqq T$, as was noted above. Q. E. D.

Now we can state the main result of the paper: $f_\epsilon \in L_p(E)$ *for all* $\epsilon > 0$, $\| f_\epsilon \|_p \leqq A_p \| f \|_p$, *where A_p is a positive constant depending only on k, n, and p, and $\| \ \|_p$ is the norm in $L_p(E)$. Also, f_ϵ converges in $L_p(E)$ as $\epsilon \to 0$.* The proof of this result will be completed in Sections 5 and 6.

3. The L_2 theory. As in the argument presented in [1] it is first necessary to derive the result $\| f_\epsilon \|_2 \leqq c \| f \|_2$, which proceeds very naturally via Fourier transform techniques. If a function $F(x,t)$ is defined on E, we shall imagine it to be extended to a function defined on R^{n+1}, identically zero for $t \leqq 0$. Then we define the Fourier transform

$$\hat{F}(x,t) = \int_{R^{n+1}} \exp[i(x\cdot\xi + t\tau)]F(\xi,\tau)\,d\xi d\tau.$$

LEMMA 2. *For all* $(x,t) \in R^{n+1}$, $|\hat{k}_{\epsilon,T}(x,t)| \leqq c$.

Proof. By definition of $\hat{k}_{\epsilon,T}$ and by (2.1)

$$(3.1)\qquad \begin{aligned} \hat{k}_{\epsilon,T}(x,t) &= \int_\epsilon^T \frac{d\tau}{\tau}\int_{R^n}\Omega(\xi)\exp[i(\phi(\tau)x\cdot\xi + t\tau)]\,d\xi \\ &= \int_\epsilon^T \frac{e^{it\tau}}{\tau}\,\hat{\Omega}(\phi(\tau)x)\,d\tau, \end{aligned}$$

where $\hat{\Omega}$ is the (n-dimensional) Fourier transform of Ω. Let $y \in R^n$. By III

$$\begin{aligned} |\hat{\Omega}(y)| &= \left|\int_{R^n}\Omega(\xi)\exp(i\xi\cdot y)\,d\xi\right| \\ &= \left|\int_{R^n}\Omega(\xi)[\exp(i\xi\cdot y) - 1]\,d\xi\right| \\ &\leqq \int_{R^n}|\Omega(\xi)|\,|\xi\cdot y|\,d\xi \\ &\leqq c|y|. \end{aligned}$$

This is a useful estimate for small $|y|$. For large $|y|$ and any $z \in R^n$

$$(3.2)\qquad \begin{aligned} \hat{\Omega}(y) &= \int_{R^n}\Omega(\xi)\exp(i\xi\cdot y)\,d\xi \\ &= \int_{R^n}\Omega(\xi+z)\exp(i\xi\cdot y + iz\cdot y)\,d\xi. \end{aligned}$$

Thus

$$(3.3)\qquad \hat{\Omega}(y)\exp(-iz\cdot y) = \int_{R^n}\Omega(\xi+z)\exp(i\xi\cdot y)\,d\xi;$$

subtracting (3.3) from (3.2),

$$\hat{\Omega}(y)[1-\exp(-iz\cdot y)] = \int_{R^n}[\Omega(\xi) - \Omega(\xi+z)]\exp(i\xi\cdot y)\,d\xi.$$

Thus, choosing $z = \pi \dfrac{y}{|y|^2}$,

$$2|\hat{\Omega}(y)| \leqq \int_{R^n}\left|\Omega(\xi) - \Omega\left(\xi + \pi\frac{y}{|y|^2}\right)\right|\,d\xi.$$

By VII

$$2|\hat{\Omega}(y)| \leqq c|y|^{-1},$$

an estimate useful for large $|y|$. Combining this with the estimate for small $|y|$,

$$|\hat{\Omega}(y)| \leqq c \min(|y|, |y|^{-1}). \tag{3.4}$$

Utilizing the estimate (3.4) in (3.1), it follows that for $x \neq 0$

$$\begin{aligned} |\hat{k}_{\epsilon,T}(x,t)| &\leqq c \int_0^\infty \frac{1}{\tau} \min(\phi(\tau)|x|, \phi(\tau)^{-1}|x|^{-1})\, d\tau \\ &= c|x| \int_{\phi(\tau)\leqq 1/|x|} \frac{\phi(\tau)}{\tau}\, d\tau \\ &\quad + \frac{c}{|x|} \int_{\phi(\tau)>1/|x|} \frac{1}{\tau\phi(\tau)}\, d\tau. \end{aligned} \tag{3.5}$$

It follows from IV that

$$\int_{\phi(\tau)\leqq 1/|x|} \frac{\phi(\tau)}{\tau} d\tau \leqq \frac{c}{|x|},$$

and

$$\int_{\phi(\tau)>1/|x|} \frac{1}{\tau\phi(\tau)} d\tau \leqq c|x|.$$

Therefore (3.5) implies

$$|\hat{k}_{\epsilon,T}(x,t)| \leqq c|x| \frac{c}{|x|} + \frac{c}{|x|} c|x| \leqq c$$

if $x \neq 0$. If $x = 0$, then $\hat{k}_{\epsilon,T}(0,t) = 0$, by III. Q.E.D.

An immediate consequence of this lemma is

THEOREM 2. *Let $f \in L_2(E)$. Then for all $\epsilon > 0$, $f_\epsilon \in L_2(E)$, and $\| f_\epsilon \|_2 \leqq c \| f \|_2$.*

Proof. Since $f_{\epsilon,T}$ is the convolution of $k_{\epsilon,T}$ and f, it follows that almost everywhere $\hat{f}_{\epsilon,T} = \hat{k}_{\epsilon,T}\hat{f}$, so that $|\hat{f}_{\epsilon,T}(x,t)| \leqq c|\hat{f}(x,t)|$ by Lemma 2. Therefore, $\|\hat{f}_{\epsilon,T}\|_2 \leqq c\|\hat{f}\|_2$. Since $\|\hat{f}\|_2 = (2\pi)^{(n+1)/2}\|f\|_2$ and likewise for $f_{\epsilon,T}$, we have $\|f_{\epsilon,T}\|_2 \leqq c\|f\|_2$, or,

$$\begin{aligned} c^2\|f\|_2^2 \geqq \|f_{\epsilon,T}\|_2^2 &= \int_0^\infty \int_{R^n} |f_{\epsilon,T}(x,t)|^2\, dxdt \\ &\geqq \int_0^T \int_{R^n} |f_{\epsilon,T}(x,t)|^2\, dxdt \\ &= \int_0^T \int_{R^n} |f_\epsilon(x,t)|^2\, dxdt, \end{aligned} \tag{3.6}$$

since f_ϵ and $f_{\epsilon,T}$ are equal for $t \leqq T$ (cf. (2.3)). As (3.6) is valid for all $T > 0$, we obtain $c^2 \| f \|_2^2 \geqq \| f_\epsilon \|_2^2$. Q.E.D.

Example. It is of interest to compute the Fourier transform of the kernel k of Section 1, associated with a parabolic differential equation. According to (1.2), $k(x,t) = \frac{1}{t^{1+n/m}} \Omega\left(\frac{x}{t^{1/m}}\right)$, where

$$\Omega(x) = (2\pi)^{-n} \int_{R^n} (iy)^\rho \exp[P(y)] \exp(ix \cdot y)\, dy,$$

which is the Fourier transform of the function $(2\pi)^{-n}(ix)^\rho \exp[P(x)]$. According to the inversion theorem for Fourier transforms,

$$\int_{R^n} \exp(-ix \cdot y) \Omega(y)\, dy = (ix)^\rho \exp[P(x)],$$

or,

$$\hat{\Omega}(x) = (-ix)^\rho \exp[P(-x)] = (ix)^\rho \exp[P(x)].$$

Therefore, by (3.1)

$$\hat{k}_{\epsilon,T}(x,t) = \int_\epsilon^T \frac{e^{it\tau}}{\tau} (ix\tau^{1/m})^\rho \exp[P(\tau^{1/m}x)]\, d\tau$$

$$= (ix)^\rho \int_\epsilon^T \exp[it\tau + \tau P(x)]\, d\tau,$$

since $\rho_1 + \cdots + \rho_n = m$ and P is homogeneous of degree m. Thus,

$$\hat{k}_{\epsilon,T}(x,t) = \frac{(ix)^\rho}{it + P(x)} \{\exp[(it + P(x))T] - \exp[(it + P(x))\epsilon]\}.$$

Since $P(x)$ has negative real part, $\hat{k}_{\epsilon,T}$ tends to a limit $\hat{k}$ as $\epsilon \to 0$, $T \to \infty$, and

$$\hat{k}(x,t) = \frac{-(ix)^\rho}{it + P(x)} = \frac{-(it^{-1/m}x)^\rho}{i + P(t^{-1/m}x)}. \tag{3.7}$$

Thus, $\hat{k}$ is a bounded rational function of $x_1, \cdots, x_n, t$, and satisfies the homogeneity relation

$$\hat{k}(\alpha^{1/m}x, \alpha t) = \hat{k}(x,t), \quad \alpha > 0.$$

4. A method of partitioning space.

The proof of the main theorem, given in Section 5, requires a fine partition of space into cells which are not cubes, but whose dimensions are closely related to the function ϕ. The word "cell" shall always mean a set of the form

$$\{(x,t) \mid x_i^0 \leqq x_i \leqq x_i^0 + r,\ i = 1, \cdots, n;\ t_1 \leqq t \leqq t_1 + s\}$$

for positive r, s and any numbers $x_1^0, \cdots, x_n^0, t_1$. Such a cell will be said to have dimensions r, s. We wish to obtain finer and finer partitions in which the ratio $\phi(s)/r$ remains bounded above and below by positive constants. A procedure for obtaining such partitions will now be described.

Suppose that R^{n+1} is partitioned into non-overlapping cells of dimensions $\phi(a)$, a. As the first step in the subdivision process, divide each cell in half in the t-direction, and let

$$k_1 = [\phi(a)/\phi(2^{-1}a)],$$

where for any positive number b, $[b]$ is the largest integer not greater than b. Now divide each of the remaining sides of each cell into k_1 equal intervals (each of length $k_1^{-1}\phi(a)$). Thus, each of the original cells of dimensions $\phi(a)$, a is divided into $2k_1^n$ new cells of dimensions $k_1^{-1}\phi(a), 2^{-1}a$. Note that since ϕ is non-decreasing, $k_1 \geqq 1$ and the process is well-defined.

In the same fashion, let

$$k_2 = [k_1^{-1}\phi(a)/\phi(2^{-2}a)]$$

and let each cell of dimensions $k_1^{-1}\phi(a), 2^{-1}a$ be divided into $2k_2^n$ new cells of dimensions $k_1^{-1}k_2^{-1}\phi(a), 2^{-2}a$. We continue this process by induction, letting

$$k_\nu = [k_1^{-1}\cdots k_{\nu-1}^{-1}\phi(a)/\phi(2^{-\nu}a)]; \tag{4.1}$$

each cell of dimensions $k_1^{-1}\cdots k_{\nu-1}^{-1}\phi(a), 2^{-\nu+1}a$ is divided into $2k_\nu^n$ new cells of dimensions $k_1^{-1}\cdots k_\nu^{-1}\phi(a), 2^{-\nu}a$.

To see that this process is well defined, we must show that $k_{\nu+1}$ is positive, assuminug $k_i \geqq 1$, $i = 1, \cdots, \nu$. But

$$\begin{aligned} k_1^{-1}\cdots k_\nu^{-1}\phi(a)/\phi(2^{-\nu-1}a) &= k_\nu^{-1}\{k_1^{-1}\cdots k_{\nu-1}^{-1}\phi(a)/\phi(2^{-\nu}a)\} \\ &\qquad \cdot \phi(2^{-\nu}a)/\phi(2^{-\nu-1}a) \\ &\geqq k_\nu^{-1}[k_1^{-1}\cdots k_{\nu-1}^{-1}\phi(a)/\phi(2^{-\nu}a)] \\ &\qquad \cdot \phi(2^{-\nu}a)/\phi(2^{-\nu-1}a) \\ &= \phi(2^{-\nu}a)/\phi(2^{-\nu-1}a) \geqq 1, \end{aligned}$$

so that $k_{\nu+1} = [k_1^{-1}\cdots k_\nu^{-1}\phi(a)/\phi(2^{-\nu-1}a)] \geqq 1$. Note also that

$$\begin{aligned} k_1^{-1}\cdots k_\nu^{-1}\phi(a)/\phi(2^{-\nu}a) &= k_\nu^{-1}\{k_1^{-1}\cdots k_{\nu-1}^{-1}\phi(a)/\phi(2^{-\nu}a)\} \\ &< k_\nu^{-1}\{[k_1^{-1}\cdots k_{\nu-1}^{-1}\phi(a)/\phi(2^{-\nu}a)] + 1\} \\ &= k_\nu^{-1}(k_\nu + 1) \leqq 2. \end{aligned} \tag{4.2}$$

But the expression on the left in (4.2) is just $r/\phi(s)$ for a cell of dimensions

r, s, where $r = k_1^{-1} \cdots k_\nu^{-1}\phi(a)$, $s = 2^{-\nu}a$. Therefore, any cell constructed in the process has dimensions r, s satisfying

$$1 \leqq \frac{r}{\phi(s)} < 2. \tag{4.3}$$

Finally, a bound is needed on k_ν. By (4.1) (with ν replaced by $\nu+1$) and (4.2) $k_{\nu+1} \leqq [2\phi(2^{-\nu}a)/\phi(2^{-\nu-1}a)]$. Therefore, V implies $k_{\nu+1} \leqq [2c] \leqq c$. Thus, k_ν is bounded by a constant depending only on the kernel $k(x,t)$, and *independent* of a, the length in the t-direction of the original cells.

Before proceeding further, the notion of rearrangement of a function must be considered. We shall use the notation of [1]. If $f \in L_p(E)$, let $f^*(\lambda)$ be a non-increasing function for $0 < \lambda < \infty$ which is equimeasurable with $|f|$. Also let

$$\beta_f(\lambda) = \frac{1}{\lambda}\int_0^\lambda f^*(\mu)\,d\mu, \quad 0 < \lambda < \infty.$$

Let $\beta^f(\lambda)$ be a function inverse to the non-increasing function β_f. For details of the properties of these functions, see [1], pp. 91-93. We now present a direct analog of Lemma 1 of Chapter I of [1].

Let $|\mathfrak{a}|$ stand for the Lebesgue measure of a measurable subset $\mathfrak{a}$ of R^{n+1}.

LEMMA 3. *Let $f \in L_p(E)$ and let $\eta > 0$. Then there is a sequence of non-overlapping cells I^j such that*

(1) *if I^j has dimensions r, s, then* $1 \leqq \dfrac{r}{\phi(s)} < 2$;

(2) $\eta \leqq |I^j|^{-1}\int_{I^j} |f| \leqq c\eta$;

(3) $|f| \leqq \eta$ *a. e. outside* $D_\eta = \bigcup_{j=1}^{\infty} I^j$;

(4) $|D_\eta| \leqq \beta^f(\eta)$;

(5) $\eta \leqq |D_\eta|^{-1}\int_{D_\eta} |f| \leqq c\eta$.

Proof. One of the properties of $\beta_f(\lambda)$ is that $\lim_{\lambda\to\infty} \beta_f(\lambda) = 0$. Therefore, $\beta_f(\lambda) < \eta$ for sufficiently large λ. Now partition R^{n+1} into cells of dimensions $\phi(a), a$, such that $\phi(a)^n a \geqq \lambda$. Over any such cell I

$$|I|^{-1}\int_I |f| \leqq |I|^{-1}\int_0^{|I|} f^*(\mu)\,d\mu = \beta_f(|I|)$$
$$= \beta_f(\phi(a)^n a) \leqq \beta_f(\lambda) < \eta.$$

Now partition these cells as described above into cells of dimensions $k_1^{-1}\phi(a)$, $2^{-1}a$, and select for the sequence $\{I^j\}$ those cells over which the average of $|f|$ is at least η. Next partition the remaining cells into cells of dimensions $k_1^{-1}k_2^{-1}\phi(a), 2^{-2}a$, selecting for the sequence $\{I^j\}$ those where the average of $|f|$ is at least η; etc. Then the left inequality of (2) is obviously satisfied. Since each I^j is contained in a cell I coming from the previous partition, and since it was proved above that $k_\nu \leqq c$, $|I| \leqq c|I^j|$; then

$$|I^j|^{-1}\int_{I^j}|f| \leqq c|I|^{-1}\int_I |f| \leqq c\eta,$$

proving the right inequality of (2). The relation (1) is an immediate consequence of (4.3).

Next, any point outside D_η belongs to arbitrarily small cells for which the average $|f|$ is less than η. Therefore, $|f| \leqq \eta$ almost everywhere outside D_η, as a result of the existence of the derivative of the indefinite integral of $|f|$ almost everywhere. The derivative of the indefinite integral of $|f|$ is taken with respect to nets formed by the cells of our partition. Although these cells do not have a parameter of regularity, the result that the above derivative is equal to $|f|$ a. e. is still valid (cf. [5], pp. 144-150).

The remaining assertions of the lemma now follow exactly as in [1], pp. 92-93. Q. E. D.

5. The main lemma. We shall now establish the analog of Lemma 2 of Chapter I of [1]. Having this lemma, the fact that the transformation $f \rightarrow f_\epsilon$ represents a bounded mapping of L_p into L_p is an almost immediate corollary, and the proof is no different from that given in [1]. However, the proof of the following lemma is quite a bit more complicated than its correspondent in [1].

LEMMA 4. *Let* $f \in L_p(E)$ *for some* p, $1 \leqq p \leqq 2$, *and let*

$$E_\eta = \{(x,t) \mid |f_\epsilon(x,t)| > \eta\}, \quad \eta > 0.$$

Let

$$|f| \wedge \eta = \min(|f|, \eta).$$

Then

$$|E_\eta| \leqq c\eta^{-2}\int_E (|f| \wedge \eta)^2\,dxdt + c\beta^f(\eta).$$

Proof. We use the construction of cells I^j guaranteed by Lemma 3. Let ϵ, η be fixed positive numbers. Define

$$h(x,t)=\begin{cases}|I^j|^{-1}\int_{I^j} f, & (x,t)\in I^j,\\ f(x,t), & (x,t)\notin D_\eta=\bigcup_{j=1}^{\infty} I^j,\end{cases}\tag{5.1}$$

and let $g=f-h$. Then

$$\begin{gathered}\int_{I^j} g=0,\\ g\equiv 0 \text{ outside } D_\eta.\end{gathered}\tag{5.2}$$

If E_1 is the set of (x,t) where $|h_\epsilon(x,t)|\geqq\frac{1}{2}\eta$ and E_2 is the set where $|g_\epsilon(x,t)|\geqq\frac{1}{2}\eta$, then clearly $E_\eta\subset E_1\cup E_2$, since $f_\epsilon=g_\epsilon+h_\epsilon$. (Clearly, both g and h are in $L_p(E)$, and g_ϵ, h_ϵ exist a.e. by Theorem 1.)

Note that

$$\begin{aligned}\int_E (|f|\wedge\eta)^2&=\int_{|f|\leqq\eta}|f|^2+\int_{|f|>\eta}\eta^2\\ &\leqq\eta^{2-p}\int_{|f|\leqq\eta}|f|^p+\eta^2|\{(x,t)\mid |f(x,t)|>\eta\}|,\end{aligned}$$

so that $|f|\wedge\eta$ is in $L_2(E)$. Outside D_η, $|f|\leqq\eta$ a.e. by Lemma 3, so that $|h|=|f|\wedge\eta$ a.e. outside D_η by (5.1). In D_η, $|h|\leqq|I^j|^{-1}\int_{I^j}|f|\leqq c\eta$ by Lemma 3. Thus,

$$\int_E |h|^2=\int_{D\eta}|h|^2+\int_{E-D\eta}|h|^2\leqq c^2\eta^2|D_\eta|+\int_E (|f|\wedge\eta)^2,\tag{5.3}$$

and $h\in L_2(E)$. Therefore Theorem 2 implies that $\int_E |h_\epsilon|^2\leqq c\int_E |h|^2$. Thus, $c\int_E |h|^2\geqq\int_{E_1}|h_\epsilon|^2\geqq\frac{1}{4}\eta^2|E_1|$, and so by (5.3) and Lemma 3(4)

$$|E_1|\leqq c\eta^{-2}\int_E (|f|\wedge\eta)^2+c\beta^f(\eta).\tag{5.4}$$

As $|E_\eta|\leqq|E_1|+|E_2|$, it remains to prove (5.4) with E_1 replaced by E_2.

First, we need to enlarge the cell I^j. Suppose that I^j has center (x^j,t^j) and dimensions r^j, s^j: thus

$$I^j=\{(x,t)\mid |x_i-x^j{}_i|\leqq\tfrac{1}{2}r^j,\ |t-t^j|\leqq\tfrac{1}{2}s^j\}.\tag{5.5}$$

Then define the cylinder

$$\mathcal{I}^j=\{(x,t)\mid |x-x^j|\leqq(\sqrt{n}/2+1)r^j,\ |t-t^j|\leqq s^j\}.\tag{5.6}$$

Clearly $I^j \subset \mathfrak{I}^j$ and $|\mathfrak{I}^j| \leqq c|I^j|$. Let $\mathfrak{D}_\eta = \bigcup_j \mathfrak{I}^j$; then $D_\eta \subset \mathfrak{D}_\eta$ and $|\mathfrak{D}_\eta| \leqq c|D_\eta|$.

Our main task will be to establish the inequality

$$\int_{E-\mathfrak{D}_\eta} |g_\epsilon| \leqq c\int_E |g| = c\int_{D_\eta} |g|. \tag{5.7}$$

Suppose for the moment that (5.7) has been proved. Then, since

$$\int_{I^j} |h| = \left|\int_{I^j} f\right| \leqq \int_{I^j} |f|,$$

it follows that

$$\int_{I^j} |g| \leqq \int_{I^j} (|f| + |h|) \leqq 2\int_{I^j} |f|;$$

therefore, (5.7) implies that

$$\int_{E-\mathfrak{D}_\eta} |g_\epsilon| \leqq c\sum_j \int_{I^j} |g| \leqq c\sum_j 2\int_{I^j} |f| = 2c\int_{D_\eta} |f|. \tag{5.8}$$

By Lemma 3(5), (5.8) implies

$$\int_{E-\mathfrak{D}_\eta} |g_\epsilon| \leqq c\eta\,|D_\eta|. \tag{5.9}$$

Now

$$\begin{aligned} |E_2| &\leqq |\mathfrak{D}_\eta| + |E_2 - \mathfrak{D}_\eta| \\ &\leqq c\,|D_\eta| + |E_2 - \mathfrak{D}_\eta|. \end{aligned} \tag{5.10}$$

But on E_2, $|g_\epsilon| \geqq \tfrac{1}{2}\eta$. Therefore, by (5.9)

$$c\eta\,|D_\eta| \geqq \int_{E_2-\mathfrak{D}_\eta} \tfrac{1}{2}\eta = \tfrac{1}{2}\eta\,|E_2 - \mathfrak{D}_\eta|,$$

so that (5.10) implies $|E_2| \leqq c\,|D_\eta| + c\,|D_\eta| \leqq c\,|D_\eta|$. By Lemma 3(4), $|E_2| \leqq c\beta^f(\eta)$, and, combining this inequality with (5.4), Lemma 4 is proved. Thus, it is sufficient to establish (5.7).

Since $g \in L_p(E)$, $g_\epsilon(x,t)$ exists a.e. by Theorem 1. Also, since $g \equiv 0$ outside D_η,

$$\begin{aligned} g_\epsilon(x,t) &= \int_{D_\eta} k_\epsilon(x-\xi, t-\tau)\,g(\xi,\tau)\,d\xi d\tau \\ &= \sum_j \int_{I^j} k_\epsilon(x-\xi, t-\tau)\,g(\xi,\tau)\,d\xi d\tau. \end{aligned}$$

Let

$$G^j{}_\epsilon(x,t) = \int_{I^j} k_\epsilon(x-\xi, t-\tau)\,g(\xi,\tau)\,d\xi d\tau, \tag{5.11}$$

so that $g_\epsilon = \sum_j G^j_\epsilon$ a. e. Then, since $E - \mathfrak{D}_\eta \subset E - \mathfrak{J}^j$,

$$\int_{E-\mathfrak{D}_\eta} | g_\epsilon | \leqq \sum_j \int_{E-\mathfrak{D}_\eta} | G^j_\epsilon | \leqq \sum_j \int_{E-\mathfrak{J}^j} | G^j_\epsilon |.$$

Therefore, to establish (5.7) it suffices to show that

$$\int_{E-\mathfrak{J}^j} | G^j_\epsilon | \leqq c \int_{I^j} | g |. \tag{5.12}$$

Since (5.7) implies the validity of the lemma, it is seen that the lemma will be proved if (5.12) is proved. Up to this point the proof of the lemma has been exactly as in [1], but the proof of (5.12), to be given now, is considerably different.

The analysis is different depending on the relation between s^j and ϵ. For simplicity we will delete the superscript j from I^j, $\mathfrak{J}^j$, r^j, s^j, and G^j_ϵ.

Case 1. $s \leqq 2\epsilon$. In this case

$$\int_{E-\mathfrak{J}} | G_\epsilon | \leqq \int_E | G_\epsilon | = J_1 + J_2, \tag{5.13}$$

where

$$J_1 = \int_{t \geqq t^j + 2\epsilon} | G_\epsilon |; \qquad J_2 = \int_{t \leqq t^j + 2\epsilon} | G_\epsilon |.$$

Now $(\xi, \tau) \in I \Rightarrow | \tau - t^j | \leqq \frac{1}{2} s \leqq \epsilon$, so that if $t \geqq t^j + 2\epsilon$, then

$$t - \tau \geqq t^j - \tau + 2\epsilon \geqq -\epsilon + 2\epsilon = \epsilon.$$

Therefore,

$$J_1 = \int_{t^j+2\epsilon}^{\infty} \int_{R^n} \left| \int_I k(x - \xi, t - \tau) g(\xi, \tau) \, d\xi d\tau \right| dx dt.$$

Since $\int_I g = 0$ by (5.2),

$$\begin{aligned} J_1 &= \int_{t^j+2\epsilon}^{\infty} \int_{R^n} \left| \int_I [k(x-\xi, t-\tau) - k(x - x^j, t - t^j)] g(\xi, \tau) d\xi d\tau \right| dx dt \\ &\leqq \int_I J_3(\xi, \tau) | g(\xi, \tau) | \, d\xi d\tau, \end{aligned} \tag{5.14}$$

where

$$J_3(\xi, \tau) = \int_{t^j+2\epsilon}^{\infty} \int_{R^n} | k(x - \xi, t - \tau) - k(x - x^j, t - t^j) | \, dx dt. \tag{5.15}$$

Next, $J_3 \leqq J_4 + J_5$, where

$$J_4(\xi, \tau) = \int_{t^j+2\epsilon}^{\infty} \int_{R^n} | k(x - \xi, t - \tau) - k(x - x^j, t - \tau) | \, dx dt,$$

$$J_5(\xi,\tau)=\int_{t^j+2\epsilon}^{\infty}\int_{R^n}|k(x-x^j,t-\tau)-k(x-x^j,t-t^j)|\,dxdt.$$

By II,

$$J_4(\xi,\tau)=\int_{t^j+2\epsilon}^{\infty}\int_{R^n}\frac{1}{(t-\tau)\phi(t-\tau)^n}\left|\Omega\left(\frac{x-\xi}{\phi(t-\tau)}\right)-\Omega\left(\frac{x-x^j}{\phi(t-\tau)}\right)\right|dxdt.$$

Making the change of variable $y=\dfrac{x-\xi}{\phi(t-\tau)}$, the integral for J_4 becomes

$$J_4(\xi,\tau)=\int_{t^j+2\epsilon}^{\infty}\frac{dt}{t-\tau}\int_{R^n}\left|\Omega(y)-\Omega\left(y+\frac{\xi-x^j}{\phi(t-\tau)}\right)\right|dy.$$

By VII and IV,

$$J_4(\xi,\tau)\leqq c\int_{t^j+2\epsilon}^{\infty}\frac{|\xi-x^j|}{(t-\tau)\phi(t-\tau)}\,dt=c\,|\xi-x^j|\int_{t^j-\tau+2\epsilon}^{\infty}\frac{d\sigma}{\sigma\phi(\sigma)}$$

$$\leqq c\,\frac{|\xi-x^j|}{\phi(t^j-\tau+2\epsilon)}.$$

Now $t^j-\tau+2\epsilon\geqq\epsilon\geqq\frac{1}{2}s$, as shown above, and $|\xi-x^j|\leqq\dfrac{\sqrt{n}}{2}r$ (all for $(\xi,\tau)\in I$), so that

$$J_4(\xi,\tau)\leqq c\,\frac{r}{\phi(\frac{1}{2}s)}\leqq c\,\frac{r}{\phi(s)},$$

by V. Finally, Lemma 3(1) implies that $J_4(\xi,\tau)\leqq c$. Next we show also that $J_5(\xi,\tau)\leqq c$. This will imply that $J_3(\xi,\tau)\leqq c$, which by (5.14) implies $J_1\leqq c\displaystyle\int_I|g|$.

Now by II,

$$J_5(\xi,\tau)=\int_{t^j+2\epsilon}^{\infty}\int_{R^n}\left|\frac{1}{(t-\tau)\phi(t-\tau)^n}\,\Omega\left(\frac{x-x^j}{\phi(t-\tau)}\right)\right.$$
$$\left.-\frac{1}{(t-t^j)\phi(t-t^j)^n}\,\Omega\left(\frac{x-x^j}{\phi(t-t^j)}\right)\right|dxdt$$
$$\leqq J_6(\xi,\tau)+J_7(\xi,\tau),$$

where

$$J_6(\xi,\tau)=\int_{t^j+2\epsilon}^{\infty}\int_{R^n}\frac{1}{(t-\tau)\phi(t-\tau)^n}\left|\Omega\left(\frac{x-x^j}{\phi(t-\tau)}\right)-\Omega\left(\frac{x-x^j}{\phi(t-t^j)}\right)\right|dxdt,$$

$$J_7(\xi,\tau)=\int_{t^j+2\epsilon}^{\infty}\int_{R^n}\left|\frac{1}{(t-\tau)\phi(t-\tau)^n}-\frac{1}{(t-t^j)\phi(t-t^j)^n}\right|$$
$$\cdot\left|\Omega\left(\frac{x-x^j}{\phi(t-t^j)}\right)\right|dxdt.$$

In the first of these two integrals the change of variable $y=\dfrac{x-x^j}{\phi(t-\tau)}$ shall be introduced, and in the second $y=\dfrac{x-x^j}{\phi(t-t^j)}$. Thus,

$$J_6(\xi,\tau)=\int_{t^j+2\epsilon}^{\infty}\frac{dt}{t-\tau}\int_{R^n}\left|\Omega(y)-\Omega\Big(\frac{\phi(t-\tau)}{\phi(t-t^j)}y\Big)\right|dy,$$

$$J_7(\xi,\tau)=\int_{t^j+2\epsilon}^{\infty}\frac{1}{t-t^j}\left|\frac{\phi(t-t^j)^n}{\phi(t-\tau)^n}\,\frac{t-t^j}{t-\tau}-1\right|dt\int_{R^n}|\Omega(y)|\,dy.$$

Now for $(\xi,\tau)\in I$ and $t\geqq t^j+2\epsilon$, we have

$$\left|\frac{t-\tau}{t-t^j}-1\right|=\frac{|\tau-t^j|}{t-t^j}\leqq\frac{\frac{1}{2}s}{2\epsilon}\leqq\tfrac{1}{2};$$

hence,

$$\tfrac{1}{2}\leqq\frac{t-\tau}{t-t^j}\leqq\tfrac{3}{2}.$$

Therefore it follows from V that the ratio $\dfrac{\phi(t-\tau)}{\phi(t-t^j)}$ is bounded above and below by positive constants depending only on the function ϕ. Therefore, VIII implies

$$\int_{R^n}\left|\Omega(y)-\Omega\left(\frac{\phi(t-\tau)}{\phi(t-t^j)}y\right)\right|dy\leqq c\left|\frac{\phi(t-\tau)}{\phi(t-t^j)}-1\right|. \tag{5.16}$$

(Note that such an inequality is trivial if $\dfrac{\phi(t-\tau)}{\phi(t-t^j)}$ is not close to 1, and is bounded away from zero.) But by VI,

$$\left|\frac{\phi(t-\tau)}{\phi(t-t^j)}-1\right|\leqq c\,\frac{|\tau-t^j|}{t-t^j}. \tag{5.17}$$

Combining (5.16) and (5.17),

$$\begin{aligned}J_6(\xi,\tau)&\leqq c\int_{t^j+2\epsilon}^{\infty}|\tau-t^j|\,\frac{1}{(t-\tau)(t-t^j)}\,dt\\&\leqq c\,|\tau-t^j|\int_{t^j+2\epsilon}^{\infty}\frac{dt}{(t-\tau)(t-t^j)}\\&=c\,\frac{|\tau-t^j|}{2\epsilon-\frac{1}{2}s}\leqq c\,\frac{\frac{1}{2}s}{s-\frac{1}{2}s}=c.\end{aligned} \tag{5.18}$$

Also

$$\left|\frac{\phi(t-t^j)^n}{\phi(t-\tau)^n}\,\frac{t-t^j}{t-\tau}-1\right| \leqq \frac{t-t^j}{t-\tau}\left|\frac{\phi(t-t^j)^n}{\phi(t-\tau)^n}-1\right| + \left|\frac{t-t^j}{t-\tau}-1\right|$$
$$\leqq c\left|\frac{\phi(t-t^j)}{\phi(t-\tau)}-1\right| + \frac{|\tau-t^j|}{t-\tau}$$
$$\leqq c\,\frac{|\tau-t^j|}{t-\tau}.$$

Therefore, this inequality and III imply

$$J_7(\xi,\tau) \leqq c\int_{t^j+2\epsilon}^{\infty} \frac{|\tau-t^j|}{(t-\tau)(t-t^j)}\,dt \leqq c,$$

by (5.18). Combining these estimates, $J_5(\xi,\tau) \leqq c$ and so

$$J_1 \leqq c\int_I |g(\xi,\tau)|\,d\xi d\tau. \tag{5.19}$$

Next, we estimate J_2. Fortunately, this estimate is quite simple. Indeed,

$$J_2 = \int_0^{t^j+2\epsilon}\int_{R^n} |G_\epsilon(x,t)|\,dxdt$$
$$\leqq \int_0^{t^j+2\epsilon}\int_{R^n}\int_I |k_\epsilon(x-\xi,t-\tau)|\,|g(\xi,\tau)|\,d\xi d\tau\, dxdt.$$

But

$$\int_0^{t^j+2\epsilon}\int_{R^n} |k_\epsilon(x-\xi,t-\tau)|\,dxdt = \int_{\tau+\epsilon}^{t^j+2\epsilon}\int_{R^n} |k(x-\xi,t-\tau)|\,dxdt$$
$$= \int_{\tau+\epsilon}^{t^j+2\epsilon} \frac{dt}{t-\tau}\int_{R^n} |\Omega(x)|\,dx$$

by (2.1), and the latter integral is bounded by $c\log\dfrac{t^j+2\epsilon-\tau}{\epsilon} \leqq c\log\dfrac{\frac{1}{2}s+2\epsilon}{\epsilon}$ $\leqq c\log 3$. Thus, $J_2 \leqq c\int_I |g(\xi,\tau)|\,d\xi d\tau$.

Combining this estimate with (5.19) and using (5.13) gives

$$\int_{E-S} |G_\epsilon| \leqq c\int_I |g|.$$

Consequently, (5.12) has been established for all cells I^j satisfying $s^j \leqq 2\epsilon$. We now treat the other cells.

Case 2. $s > 2\epsilon$. Using again the fact that $\int_I g = 0$ (see (5.2)), we have by definition of G_ϵ (see (5.11))

$$
\begin{aligned}
\int_{E-\mathcal{A}} | G_\epsilon | &= \int_{E-\mathcal{A}} \left| \int_I k_\epsilon(x-\xi, t-\tau) g(\xi,\tau)\, d\xi d\tau \right| dxdt \\
&= \int_{\substack{E-\mathcal{A} \\ t \leqq t^j + s}} \left| \int_I k_\epsilon(x-\xi, t-\tau) g(\xi,\tau)\, d\xi d\tau \right| dxdt \\
&\qquad + \int_{\substack{E-\mathcal{A} \\ t > t^j + s}} \left| \int_I [k_\epsilon(x-\xi, t-\tau) - k_\epsilon(x-x^j, t-t^j)] \right. \\
&\qquad\qquad \left. \cdot\, g(\xi,\tau)\, d\xi d\tau \right| dxdt \\
&\leqq \int_I [J_8(\xi,\tau) + J_9(\xi,\tau)] \, | g(\xi,\tau) | \, d\xi d\tau,
\end{aligned}
\tag{5.20}
$$

where

$$J_8(\xi,\tau) = \int_{\substack{E-\mathcal{A} \\ t \leqq t^j + s}} | k_\epsilon(x-\xi, t-\tau) | \, dxdt,$$

$$J_9(\xi,\tau) = \int_{\substack{E-\mathcal{A} \\ t > t^j + s}} | k_\epsilon(x-\xi, t-\tau) - k_\epsilon(x-x^j, t-t^j) | \, dxdt.$$

The integral J_8 is easily treated. Since $k_\epsilon(x-\xi, t-\tau) = 0$ for $t < \tau + \epsilon$, and since $(\xi,\tau) \in I \Rightarrow \tau + \epsilon \leqq t^j + \frac{1}{2}s + \epsilon < t^j + s$,

$$J_8(\xi,\tau) = \int_{\tau+\epsilon}^{t^j+s} \int_{|x-x^j| > (\sqrt{n}/2+1)r} | k(x-\xi,\ t-\tau) | \, dxdt;$$

we have used the definition of $\mathcal{A}$ ((5.6)). For $(\xi,\tau) \in I$, $| \xi_i - x^j_i | \leqq \frac{1}{2}r$ according to (5.5), so that $| \xi - x^j | \leqq \frac{1}{2}\sqrt{n}\, r$. Therefore, by the triangle inequality, $| x - x^j | > (\sqrt{n}/2 + 1)r \Rightarrow | x - \xi | > r$. Hence,

$$
\begin{aligned}
J_8(\xi,\tau) &\leqq \int_{\tau+\epsilon}^{t^j+s} \int_{|x-\xi|>r} | k(x-\xi, t-\tau) | \, dxdt \\
&= \int_{\tau+\epsilon}^{t^j+s} \frac{dt}{t-\tau} \int_{|y| > r/\phi(t-\tau)} | \Omega(y) | \, dy,
\end{aligned}
$$

taking (2.1) into account. But then IX implies

$$
\begin{aligned}
J_8(\xi,\tau) &\leqq c \int_{\tau+\epsilon}^{t^j+s} (\phi(t-\tau)/r)^n \, \frac{dt}{t-\tau} \\
&= c \int_{\epsilon}^{t^j-\tau+s} (\phi(t)/r)^n \frac{dt}{t} \\
&\leqq c \int_0^{\frac{3}{2}s} (\phi(t)/r)^n \frac{dt}{t},
\end{aligned}
$$

since for $(\xi,\tau) \in \mathrm{I}$, $t^j - \tau \leqq \frac{1}{2}s$. As $\phi(t)$ is nondecreasing, and by IV, V,

$$J_8(\xi,\tau) \leqq c\phi(\tfrac{3}{2}s)^{n-1}r^{-n}\int_0^{\frac{3}{2}s}\frac{\phi(t)}{t}\,dt$$
$$\leqq c\phi(\tfrac{3}{2}s)^n r^{-n} \leqq c(\phi(s)/r)^n.$$

Thus, Lemma 3 (1) implies $J_8(\xi,\tau) \leqq c$.

Finally, we show that $J_9(\xi,\tau) \leqq c$. Once this is established, (5.20) will imply $\int_{E-I} | G_\epsilon | \leqq c \int_I | g(\xi,\tau) | \, d\xi d\tau$. Therefore, (5.12) will have been established for all cells I^j satisfying $s^j > 2\epsilon$. As (5.12) was proved for $s^j \leqq 2\epsilon$ under the first case, (5.12), and hence (5.7), will have been completely proved. As mentioned above, the proof of the lemma will then have been completed.

Now

$$(5.21) \qquad J_9(\xi,\tau) \leqq \int_{t^j+s}^{\infty}\int_{R^n} | k(x-\xi, t-\tau) - k(x-x^j, t-t^j) | \, dx dt,$$

since $t-\tau = (t-t^j) + (t^j-\tau) \geqq s - \frac{1}{2}s = \frac{1}{2}s > \epsilon$. (Recall that $(\xi,\tau) \in I \Rightarrow | t^j - \tau | \leqq \frac{1}{2}s$.) The bound for $J_9(\xi,\tau)$ in (5.21) is quite similar to the expression (5.15) for $J_3(\xi,\tau)$, except that 2ϵ is replaced by s. The fact that $J_9(\xi,\tau) \leqq c$ is proved in exactly the same way as the estimate $J_3(\xi,\tau) \leqq c$ is derived, except that the proof for $J_9(\xi,\tau)$ is somewhat simpler at some stages. Thus, $J_9(\xi,\tau) \leqq c$. By the remarks made in the previous paragraph, Lemma 4 has now been established. Q. E. D.

6. Convergence in L_p.

As mentioned in the first few lines of Section 5, the following theorem can now be established.

THEOREM 3. *Let $1 < p < \infty$, and let $f \in L_p(E)$. Then f_ϵ, defined by*

$$f_\epsilon(x) = \int_E k_\epsilon(x-y) f(y)\, dy,$$

belongs to $L_p(E)$, and there exists a constant A_p, depending only on p and the kernel k, such that

$$\| f_\epsilon \|_p \leqq A_p \| f \|_p.$$

The proof is exactly the same as given in [1], pp. 97-99, using Lemma 4.

Finally, we establish the existence of $\lim_{\epsilon\to 0} f_\epsilon$ in the space L_p.

THEOREM 4. *Let $1 < p < \infty$, and let $f \in L_p(E)$. Then f_ϵ converges in $L_p(E)$ as $\epsilon \to 0$.*

Proof. Suppose $F(x,t)$ is continuously differentiable for all $(x,t) \in R^{n+1}$, has compact support, and vanishes for $t \leqq 0$. Then

$$
\begin{aligned}
F_\epsilon(x,t) &= \int_0^\infty \int_{R^n} k_\epsilon(x-\xi, t-\tau) F(\xi,\tau)\, d\xi d\tau \\
&= \int_0^{\max(0,t-\epsilon)} \int_{R^n} k(x-\xi, t-\tau) F(\xi,\tau)\, d\xi d\tau \qquad (6.1)\\
&= \int_0^{\max(0,t-\epsilon)} \int_{R^n} k(x-\xi, t-\tau)\, [F(\xi,\tau) - F(x,t)]\, d\xi d\tau,
\end{aligned}
$$

by III. By the properties of F, $|F(\xi,\tau) - F(x,t)| \leqq M[|x-\xi| + |t-\tau|]$ for some constant M. Thus, if $0 < \epsilon < \delta$,

$$
\begin{aligned}
|F_\epsilon(x,t) - F_\delta(x,t)| &= \Big| \int_{\max(0,t-\delta)}^{\max(0,t-\epsilon)} \int_{R^n} k(x-\xi, t-\tau) \\
&\qquad \cdot [F(\xi,\tau) - F(x,t)]\, d\xi d\tau \Big| \qquad (6.2)\\
&\leqq M \int_{t-\delta}^{t} \int_{R^n} |k(x-\xi,\, t-\tau)|\, [|x-\xi| + t-\tau]\, d\xi d\tau \\
&= M \int_{t-\delta}^{t} \frac{d\tau}{t-\tau} \int_{R^n} |\Omega(y)|\, [|\phi(t-\tau) y| + t-\tau]\, dy,
\end{aligned}
$$

by (2.1). By III, IV,

$$
\begin{aligned}
|F_\epsilon(x,t) - F_\delta(x,t)| &\leqq M \int_{t-\delta}^{t} \frac{\phi(t-\tau)}{t-\tau}\, d\tau \int_{R^n} |y|\, |\Omega(y)|\, dy \\
&\quad + M \int_{t-\delta}^{t} d\tau \int_{R^n} |\Omega(y)|\, dy \\
&\leqq Mc \int_0^\delta \frac{\phi(\sigma)}{\sigma}\, d\sigma + Mc\delta \\
&\leqq Mc\phi(\delta) + Mc\delta.
\end{aligned}
$$

Therefore $F_\epsilon(x,t)$ converges as $\epsilon \to 0$, uniformly for $x \in R^n$, $t \geqq 0$. Denote the limit function by $F_0(x,t)$. By (6.1) and (2.2) $F_\epsilon(x,t) \equiv F_0(x,t)$ for $\epsilon < 1$ and for sufficiently large t depending on the support of F. Thus, if $F(x,t)$ vanishes for $t \geqq T$, and if $\epsilon < 1$,

$$
\begin{aligned}
\| F_\epsilon - F_0 \|_p{}^p &= \int_0^{T+1} \int_{R^n} |F_\epsilon(x,t) - F_0(x,t)|^p\, dx dt \\
&= \int_0^{T+1} \int_{|x| \leqq A} + \int_0^{T+1} \int_{|x| > A} |F_\epsilon - F_0|^p\, dx dt,
\end{aligned}
$$

for any positive A. Since F_ϵ converges uniformly to F_0,

$$(6.3)\qquad \limsup_{\epsilon\to 0}\|F_\epsilon - F_0\|_p{}^p = \limsup_{\epsilon\to 0}\int_0^{T+1}\int_{|x|>A}|F_\epsilon - F_0|^p\,dx\,dt.$$

Letting $\epsilon\to 0$ in (6.2) and then replacing δ by ϵ in the resulting expression gives

$$(6.4)\qquad \begin{aligned} |F_0(x,t) - F_\epsilon(x,t)| &= \Big|\int_{\max(0,t-\epsilon)}^{t}\int_{R^n} k(x-\xi,t-\tau) \\ &\qquad\cdot[F(\xi,\tau) - F(x,t)]\,d\xi d\tau\Big| \\ &\leqq \int_{t-\epsilon}^{t}\int_{R^n}|k(x-\xi,t-\tau)|\,|F(\xi,\tau) - F(x,t)|\,d\xi d\tau. \end{aligned}$$

Now suppose that $F(x,t)\equiv 0$ for $|x|\geqq B$. Then for $|x|>A>2B$, (6.4) becomes

$$\begin{aligned} |F_0(x,t) - F_\epsilon(x,t)| &\leqq \int_{t-\epsilon}^{t}\int_{|\xi|<B}|k(x-\xi,t-\tau)|\,|F(\xi,\tau)|\,d\xi d\tau \\ &\leqq \max|F|\int_{t-\epsilon}^{t}\int_{|x-y|<B}|k(y,t-\tau)|\,dy d\tau \\ &= \max|F|\int_{t-\epsilon}^{t}\frac{d\tau}{t-\tau}\int_{|x-\phi(t-\tau)y|<B}|\Omega(y)|\,dy, \end{aligned}$$

by (2.1). Therefore, if $\gamma = \max|F|$,

$$(6.5)\qquad |F_0(x,t) - F_\epsilon(x,t)| \leqq \gamma\int_0^{\epsilon}\frac{d\tau}{\tau}\int_{|x-\phi(\tau)y|<B}|\Omega(y)|\,dy.$$

Now for $|x|>A>2B$, $|x-\phi(\tau)y|<B$ implies

$$\phi(\tau)|y| \geqq |x| - |x-\phi(\tau)y| > |x| - B > |x| - \tfrac{1}{2}|x| = \tfrac{1}{2}|x|,$$

so that (6.5) implies

$$|F_0(x,t) - F_\epsilon(x,t)| \leqq \gamma\int_0^{\epsilon}\frac{d\tau}{\tau}\int_{|y|>|x|/2\phi(\tau)}|\Omega(y)|\,dy,\quad |x|>A.$$

Taking IX and IV into account, we have for $|x|>A$,

$$\begin{aligned} |F_0(x,t) - F_\epsilon(x,t)| &\leqq \gamma c\int_0^{\epsilon}\frac{1}{\tau}\,(2\phi(\tau)/|x|)^n\,d\tau \\ &\leqq \gamma c\phi(\epsilon)^{n-1}|x|^{-n}\int_0^{\epsilon}\frac{\phi(\tau)}{\tau}\,d\tau \\ &\leqq \gamma c\phi(\epsilon)^n|x|^{-n}. \end{aligned}$$

Therefore, the integral in (6.3) may be estimated as follows:

$$\int_0^{T+1}\int_{|x|>A} |F_\epsilon - F_0|^p \,dx\,dt \leqq (T+1)\gamma^p c^p \phi(\epsilon)^{np} \int_{|x|>A} |x|^{-np}\,dx$$
$$\leqq (T+1)\gamma^p c\phi(\epsilon)^{np} A^{-n(p-1)},$$

as $p > 1$. Thus, this integral tends to zero as $\epsilon \to 0$, so that (6.3) implies

$$\lim_{\epsilon\to 0} \| F_\epsilon - F_0 \|_p = 0$$

for all functions F of the class considered.

Since the transformations mappings f into f_ϵ are uniformly bounded by Theorem 3, and since f_ϵ converges for f in a dense subset of $L_p(E)$, it follows immediately that f_ϵ converges for all $f \in L_p(E)$. Q. E. D.

7. **Remarks.** A few remarks on the special case of Section 1 are in order. In the first place, the kernel k has a Fourier transform $\hat{k}$ which is a bounded rational function of $x_1, \cdots, x_n, t$, as shown at the end of Section 3; see in particular (3.7). Our results show that $\hat{k}$ is a multiplier in $L_p(E)$ in the sense that if $f \in L_2(E) \cap L_p(E)$, then the inverse Fourier transform f_0 of $\hat{k}\hat{f}$ is also in $L_p(E)$, and $\| f_0 \|_p \leqq A_p \| f \|_p$.

There are several theorems in the paper of Calderón and Zygmund [1] which are consequences of their Lemma 2 which corresponds to our Lemma 4, and analogs of these theorems could also have been stated in the present paper. However, we have not been able to obtain several very important results found in [1]. For example, it would be fruitful to know that f_ϵ converges pointwise a. e. as well as in mean. Also, one would like to know that the function

$$\tilde{f}_0(x,t) = \sup_{\epsilon>0} | f_\epsilon(x,t) |$$

is in $L_p(E)$ whenever f is and that $\| \tilde{f}_0 \|_p \leqq A_p \| f \|_p$.

REFERENCES.

[1] A. P. Calderón and A. Zygmund, "On the existence of certain singular integrals," *Acta Mathematica*, vol. 88 (1952), pp. 85-139.
[2] B. Frank Jones, Jr., "Singular integrals and parabolic equations," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 501-503.
[3] W. Pogorzelski, "Étude de la solution fondamentale de l'équation parabolique," *Ricerche Matematica*, vol. 5 (1956), pp. 25-57.
[4] Paul C. Rosenbloom, "Linear partial differential equations," *Numerical Analysis and Partial Differential Equations*, vol. 5 (1958), pp. 45-195.
[5] A. C. Zaanen, *An Introduction to the Theory of Integration*, North-Holland, 1958.
[6] A. Zygmund, *Trigonometric Series*, I, Cambridge University Press, 1959.

ERRATA.

Errata to F. J. Hahn,[1] "Affine transformations of compact Abelian groups," *American Journal of Mathematics*, vol. 85 (1963), pp. 428-446.

It was pointed out to the author by W. Parry that Lemma 4 on page 430 is incorrect as it stands. However if we make the assumption that $\hat{X}$ has no elements of finite order the lemma is correct. The known proofs of this fact are radically different from the approach used in the paper. The proof I give here is due to S. Kakutani. With his permission I give it rather than my original correction because of its elegance and ingenuity.

On p. 430 replace Lemma 4 and its proof by the following:

LEMMA 4. *Let* $S = (x_0, A) \in \mathrm{Aff}(X)$ *and suppose* S *is ergodic. It follows that if* ξ *is a character of* X *and if there is a positive integer* n *such that* $V_S{}^n\xi = \xi$ *then* ξ *is of finite order.*

Proof. We suppose ξ is a character of X and $V_S{}^n\xi = \xi$ for some positive integer n. We may assume n is the smallest such integer. To show that ξ is of finite order it suffices to show that the range of ξ is finite. The idea of the proof is to define a particular continuous homomorphism $\psi \colon X \to T^m$ (T^m is the m dimensional torus) and to let K be the closed subgroup $\psi(X)$. We let μ and θ be normalized Haar measure on X and K respectively. We then define a continuous, one to one map $\sigma \colon K \to K$ such that $\psi S = \sigma\psi$, moreover we show that $\sigma^n =$ identity. We use the ergodicity of S to show that σ is ergodic. Since σ is also periodic it follows that K is finite. We use this fact to show ξ is of finite order. We now fill in the details of the argument.

Since $V_S\xi = \xi(x_0)V_A\xi$ we see that $V_S{}^n\xi = \xi$ implies $V_A{}^n\xi = \xi$. There is a smallest positive integer m such that $V_A{}^m\xi = \xi$ and moreover $n = md$ where d is a positive integer. We observe $V_S{}^m\xi = c\xi$ consequently $c^d\xi = V_S{}^{md}\xi = V_S{}^n\xi = \xi$ and thus $c^d = 1$. For later reference we note

$$c = \xi(x_0 + Ax_0 + \cdots + A^{m-1}x_0).$$

We define $\psi \colon X \to T^m$ by $\psi(x) = (\xi(x), \xi(Ax), \xi(A^2x), \cdots, \xi(A^{m-1}x))$. We define σ first on T^m and then restrict it to K. We let

$$\sigma(v_1, v_2, \cdots, v_m) = (\xi(x_0)v_2, \xi(Ax_0)v_3, \cdots, \xi(A^{m-2}x_0)v_m, \xi(A^{m-1}x_0)v_1).$$

---

[1] The author was partially supported by NSF G-25222.

Since $c = \xi(x_0 + Ax_0 + \cdots + A^{n-1}x_0)$ and $c^d = 1$ we have $\sigma^m(v_1, v_2, \cdots, v_m) = (cv_1, cv_2, \cdots, cv_m)$ and $\sigma^n = \sigma^{md} =$ identity. Thus σ is periodic. We compute easily that $\psi S = \sigma\psi$ so $\sigma\colon K \rightarrow K$.

We now wish to show that $\sigma\colon K \rightarrow K$ is measure preserving. For each Borel set $E \subset K$ define $\theta(E) = \mu(\psi^{-1}(E))$. We see that $\theta(K) = 1$, $\theta(vE) = \theta(E)$ and if E is open then $\theta(E) > 0$. This shows that θ is Haar measure on K. The σ invariance of θ follows from

$$\theta(\sigma E) = \mu(\psi^{-1}\sigma E) = \mu(S\psi^{-1}E) = \mu(\psi^{-1}E) = \theta(E).$$

If $\theta(E) > 0$ we see that $\mu(\psi^{-1}E) > 0$ and thus the ergodicity of (X, S, μ) implies that (K, σ, θ) is ergodic. Thus σ is simultaneously ergodic and periodic with period n. For each point $k \in K$ we must have $\theta(k) > 0$. If this were not true then there would be an open set U such that $0 < \theta(U) < 1/2n$. Let $W = \bigcup_{j=0}^{n-1} \sigma^j U$. We see that $0 < \theta(W) \leqq \frac{1}{2}$ and W is σ invariant. This cannot be because σ is ergodic. The same type agrument shows $k \in K$ implies $\theta(k) = 1/n$. Thus K is finite. Since

$$K = \{(\xi(x), \xi(Ax), \cdots, \xi(A^{n-1}x)) : x \in X\}$$

it follows that ξ has finite range and is thus of finite order.

Also on page 430 following the last line insert the following sentence in the hypothesis of Theorem 4: *Moreover suppose that $\hat{X}$ has no elements of finite order.*

In connection with this error I remark that Lemma 4 and Theorem 4 are not used in Section 3. Sections 2, 4 and 5 concern themselves only with the case where X is a torus so $\hat{X}$ has no elements of finite order. Thus the corrections have no effect on the remainder of the paper.

YALE UNIVERSITY.

ADDENDUM.

Addendum to Morris Newman, "Normal congruence subgroups of the modular group," *American Journal of Mathematics,* vol. 85 (1963), pp. 419-427.

Addendum. In the proof of Theorem 2, the argument at the end of page 426 requires a special dicission in the case $m = pq$, p, q distinct primes. Choosing $d = p$, $d = q$ we find that

$$(\Gamma(n) : \Gamma(np)) = (\Gamma(n) : \Gamma(nq))$$

which implies that $p^3 = q^3$, $p^3 = q(q^2 - 1)$, $p(p^2 - 1) = q^3$ or $p(p^2 - 1) = q(q^2 - 1)$. None of these cases is possible and the proof is concluded as in the paper.

A NEW GROUP ALGEBRA FOR LOCALLY COMPACT GROUPS.

By John Ernest.

When an abstract group is represented by linear transformations on a vector space, the group of transformations generates an algebra. In the case of finite dimensional representations of a finite group, the algebra consists of linear combinations of elements in the range of the representation. It is therefore natural to associate an algebra with the abstract group in a similar way be defining the group algebra. The representation theory of the group is then equivalent to the representation theory of an algebra.

Unfortunately there appears to be no single natural analogue of the concept of group algebra, to be used when studying infinite dimensional representations of topological groups. In the case of unitary representations of locally compact groups, the most popular choice for the role of "group algebra" has been $L^1(G)$. The crucial sense in which $L^1(G)$ generalizes the concept of the group algebra of a finite group is that the strongly continuous unitary representation theory of G is equivalent to the no-where trivial $*$-representation theory of $L^1(G)$.

One sense in which the analogue breaks down is that $L^1(G)$ is not the same "kind" of algebra as that associated with the range of a representation. This deficiency has been ameliorated by the following procedure. If we renorm $L^1(G)$ by defining $\| f \|_{C^*} = \operatorname*{Sup}_{T} \| T_f \|$, where the supremum is taken over all no-where trivial $*$-representations T of $L^1(G)$, and then complete $L^1(G)$ relative to this norm, we obtain a C^*-algebra, $C^*(G)$, which is called the group C^*-algebra. Once again we have that the strongly continuous unitary representation theory of G is equivalent to the no-where trivial $*$-representation theory of $C^*(G)$. (Cf. [6] and [7].) Unforunately the sense in which $L^1(G)$ and $C^*(G)$ are generated by the group G is not analogous to the group algebra of a finite group. Except when G is discrete, $L^1(G)$ and $C^*(G)$ do not even contain the group.

Another candidate for the role of group algebra, which does contain the group, is the measure group algebra $M(G)$ of all complex valued finite regular measures on G, where multiplication is convolution. A group element x

---

Received September 23, 1963.

appears in $M(G)$ as a measure μ_x defined, for any subset B of G, by $\mu_x(B) = 0$ if $x \notin B$ and $\mu_x(B) = 1$ if $x \in B$. The study of $M(G)$ for non-abelian locally compact groups has been largely abortive, since the crucial requirement for a group algebra breaks down. The representation theories of G and $M(G)$ are *not* equivalent. The norm topology of $M(G)$ induces the discrete topology on G. Thus while every *-representation of $M(G)$ gives rise, by restriction, to a representation of G by unitary operators, this representation need not be strongly continuous relative to the original topology of G.

Unlike the finite dimensional case, there is more than one way to associate an algebra with the range of an infinite dimensional unitary representation L of a locally compact group G. Experience indicates that the most useful algebra to consider is the weakly closed *-algebra of operators $\mathfrak{A}(L)$, generated by the range of the representation L. Since such a von Neumann algebra can be associated with every representation, one asks, by analogy with the group algebra of a finite group, whether such an algebra can be associated with the original locally compact group? From this point of view, the "flaw" in $M(G)$ is that there is no other topology, in addition to the norm topology. We shall see later that this is where the topology of G was lost.

The purpose of this paper is to introduce another candidate for the title "group algebra" of a locally compact group. This algebra, which we call the *big group algebra* and which we denote $\mathfrak{A}(G)$, is defined at the beginning of Section 1. $\mathfrak{A}(G)$ is a C^*-algebra in which analogues of the operator topologies are defined. These are called the weak, strong, σ-weak and σ-strong topologies of $\mathfrak{A}(G)$. The adjective "big" derives from the fact that each of the previously mentioned candidates, $L^1(G)$, $C^*(G)$, and $M(G)$, are canonically embedded, one-to-one and norm decreasing, in $\mathfrak{A}(G)$. These classical group algebras are dense in $\mathfrak{A}(G)$, relative to each of the topologies weak, σ-weak, strong and σ-strong. The group G is canonically embedded in $\mathfrak{A}(G)$ as a group of unitary elements. The topology of G is obtained as the restriction to G of any one of the topologies weak, σ-weak, strong, and σ-strong of $\mathfrak{A}(G)$. G is discrete in the norm topology of $\mathfrak{A}(G)$. The group G generates $\mathfrak{A}(G)$ in the sense that the subalgebra of finite linear combinations of elements of G, is dense in $\mathfrak{A}(G)$, relative to any one of the four topologies, weak, σ-weak, strong and σ-strong.

The algebra $\mathfrak{A}(G)$ may be represented isomorphically as a von Neumann algebra acting on an (in general) non-separable Hilbert space. This isomorphic representation of $\mathfrak{A}(G)$ as a von Neumann algebra is a homeomorphism relative to the σ-weak and σ-strong topologies.

The algebra $\mathfrak{A}(G)$ may be identified, as a Banach space, with the second conjugate space of the group C^*-algebra, $C^*(G)$. Under this identification, the weak topology of $\mathfrak{A}(G)$ corresponds to the weak*-topology of the second conjugate space of $C^*(G)$.

Finally, the separable strongly continuous unitary representation theory of G is completely equivalent to the separable normal *-representation theory of $\mathfrak{A}(G)$. Every separable strongly continuous unitary representation L of G may be extended to a normal *-representation L' of $\mathfrak{A}(G)$ onto $\mathfrak{A}(L)$. This extension is unique. Indeed the restriction of a separable normal *-representation T of $\mathfrak{A}(G)$, to the group G, gives a strongly continuous unitary representation L of G such that $T = L'$.

This theory had its beginnings when the author was an NSF postdoctoral fellow in residence at the Institute for Advanced Study. In addition to expressing his gratitude to his home institution, the University of Rochester, the author wishes to thank the Crandall Library of Glens Falls, New York, for the use of their facilities during a portion of this research. This work was supported in part by grants from the National Science Foundation, NSF-G23799 and NSF-GP1620.

1. Definition of the big group algebra. Let G denote a separable locally compact group. Let $\mathfrak{H}$ denote an infinite dimensional separable complex Hilbert space. Let G^o denote the collection of strongly continuous unitary representations of G, with representation space $\mathfrak{H}$. Let $\mathfrak{A}(G)$ denote the set of all maps J on G^o, whose values are operators on the space $\mathfrak{H}$, satisfying the following two properties.

(1) $\operatorname{Sup}\{\| J(L) \| : L \in G^o\} < +\infty$

(2) If $M, N \in G^o$ and if U is an isometric mapping of the representation space of $M \oplus N$ onto $\mathfrak{H}$, then

$$J(U(M \oplus N)U^{-1}) = U(J(M) \oplus J(N))U^{-1}.$$

The elements of $\mathfrak{A}(G)$ will be called *options*. We define a *-algebra structure for $\mathfrak{A}(G)$ as follows. If J, K are options in $\mathfrak{A}(G)$ and α, β are complex numbers, then

(i) $\| J \| = \operatorname{Sup}\{\| J(L) \| : L \in G^o\}$

(ii) $(\alpha J + \beta K)(L) = \alpha J(L) + \beta K(L)$, for all L in G^o

(iii) $(JK)(L) = J(L) \cdot K(L)$, for all L in G^o

(iv) $J^*(L) = (J(L))^*$, for all L in G^o.

Definition 1.1. The collection of options $\mathcal{A}(G)$, with this *-algebra structure, will be called the *big group algebra.*

THEOREM 1.2. *$\mathcal{A}(G)$ is a C^*-algebra.*

Proof. A routine verification gives that $\mathcal{A}(G)$ is closed under the *-algebra operations. It is easy to verify that if α, β are complex numbers, and J and K are in $\mathcal{A}(G)$, then $\| \alpha J + \beta K \| \leqq | \alpha | \, \| J \| + | \beta | \, \| K \|$, $\| JK \| \leqq \| J \| \cdot \| K \|$, $\| J^* \| = \| J \|$ and $\| J^* J \| = \| J \|^2$. If $\{J_i\}$ is a Cauchy sequence in $\mathcal{A}(G)$, then, for each L in G^c, $\{J_i(L)\}$ is a Cauchy sequence of operators which converges to an operator $J(L)$. A simple verification now gives that the map J defined in this way is an element of $\mathcal{A}(G)$ and that the sequence $\{J_i\}$ converges to J.

Definition 1.3. The *weak topology* of $\mathcal{A}(G)$ is defined to be the smallest topology such that the functions $J \rightarrow (J(L)\psi, \phi)$ are continuous, for all L in G^c and ψ, ϕ in $\mathfrak{H}$.

A neighborhood system for a point J in $\mathcal{A}(G)$ consists of the collection of sets of the form

$$\begin{aligned} &U(J : L^1, \cdots, L^n, \psi_1, \cdots, \psi_n, \phi_1, \cdots, \phi_n) \\ &\qquad = \{K : K \in \mathcal{A}(G) \text{ and } | ((J(L^i) - K(L^i))\psi_i, \phi_i) | \leqq 1, i = 1, \cdots, n\} \end{aligned}$$

where L^i is in G^c and ψ_i and ϕ_i are in $\mathfrak{H}$, for $i = 1, \cdots, n$.

The weak topology is Hausdorff and locally convex. The linear operations in $\mathcal{A}(G)$ are weakly continuous and in addition, if K is a fixed element of $\mathcal{A}(G)$, the three maps $J \rightarrow KJ$, $J \rightarrow JK$ and $J \rightarrow J^*$, of $\mathcal{A}(G)$ into $\mathcal{A}(G)$, are each weakly continuous.

2. $\mathcal{A}(G)$ contains the group and the classical group algebras.

In this section we consider the three classical group algebras, $L^1(G)$, $C^*(G)$ and $M(G)$, referred to in the introduction. We first exhibit a canonical isomorphic embedding of each of these group algebras into $\mathcal{A}(G)$.

Definition 2.1. For each ν in $M(G)$, define $\hat{\nu}$ to be the operator valued function on G^c defined by the formula

$$(\hat{\nu}(L)\psi, \phi) = \int_G (L_x\psi, \phi)\, d\nu(x)$$

for all L in G^c and all ψ, ϕ in $\mathfrak{H}$.

Since $L^1(G)$ may be identified with those measures in $M(G)$ which are absolutely continuous with respect to Haar measure, this defines an embedding of $L^1(G)$ into $\mathfrak{A}(G)$. If $f \in L^1(G)$, $\hat{f}$ is the operator valued function on G^o defined by the requirement

$$(\hat{f}(L)\psi, \phi) = \int_G f(x)(L_x\psi, \phi)\, d\mu(x)$$

for all ψ, ϕ in $\mathfrak{H}$. In this formula, μ denotes Haar measure.

THEOREM 2.2. *The map $\nu \to \hat{\nu}$ and $f \to \hat{f}$ defined above is a norm decreasing isomorphism of $M(G)$ and $L^1(G)$, into $\mathfrak{A}(G)$. The map of $L^1(G)$ may be extended uniquely to an isometric isomorphism of $C^*(G)$ into $\mathfrak{A}(G)$.*

Proof. We first verify that $\nu \to \hat{\nu}$ is an isomorphism of $M(G)$ into $\mathfrak{A}(G)$. Suppose ν is an element of $M(G)$. The norm in $M(G)$ is defined to be $\|\nu\| = |\nu|(G)$, where $|\nu|$ is the measure on G defined by

$$|\nu|(E) = \operatorname{Sup} \sum |\nu(E_i)|,$$

where the supremum is taken over all finite collections of pairwise disjoint Borel sets E_i, whose union is E.

If f is a complex valued ν-integrable function on G, then

$$\left|\int_G f\, d\nu\right| \leqq \|\nu\| \cdot \operatorname{Sup}\{|f(x)| : x \in G\}.$$

If $L \in G^o$ and $\psi \in \mathfrak{H}$, $\|\psi\| = 1$, then

$$\begin{aligned}\|\hat{\nu}(L)\psi\|^2 &= (\hat{\nu}(L)\psi, \hat{\nu}(L)\psi)\\ &= \int_G (L_x\psi, \hat{\nu}(L)\psi)\, d\nu(x)\\ &\leqq \|\nu\| \cdot \operatorname{Sup}\{|(L_x\psi, \hat{\nu}(L)\psi)| : x \in G\}\\ &\leqq \|\nu\| \cdot \operatorname{Sup}\{\|L_x\psi\| \cdot \|\hat{\nu}(L)\psi\| : x \in G\}\\ &\leqq \|\nu\| \cdot \|\hat{\nu}(L)\psi\|.\end{aligned}$$

Hence $\|\hat{\nu}\| = \operatorname{Sup}\{\|\hat{\nu}(L)\psi\| : L \in G^o, \psi \in \mathfrak{H}, \|\psi\| = 1\} \leqq \|\nu\|$.

Thus we have verified that $\hat{\nu}$ satisfies condition (1) in the definition of $\mathfrak{A}(G)$ and that the embedding is norm-decreasing. We leave to the reader the verification that $\hat{\nu}$ also satisfies condition (2) of definition 1.1.

The reader may verify that the mapping $\nu \to \hat{\nu}$ preserves the linear structure. We apply the Fubini theorem to show that the multiplicative structure is preserved. Suppose $\nu = \nu_1 * \nu_2$, where ν, ν_1, ν_2 are in $M(G)$. Then for all L in G^o and ψ, ϕ in $\mathfrak{H}$,

$$\begin{aligned}(\hat{\nu}(L)\psi,\phi) &= \int_G (L_x\psi,\phi)\,d(\nu_1 * \nu_2)(x)\\ &= \int_{G\times G} (L_{xy}\psi,\phi)\,d(\nu_1\times\nu_2)(x,y)\\ &= \int_G \int_G (L_xL_y\psi,\phi)\,d\nu_2(y)\,d\nu_1(x)\\ &= \int_G (\hat{\nu}_1(L)L_y\psi,\phi)\,d\nu_2(y)\\ &= (\hat{\nu}_2(L)\psi,\hat{\nu}_1(L)^*\phi)\\ &= ((\hat{\nu}_1\cdot\hat{\nu}_2)(L)\psi,\phi).\end{aligned}$$

Thus the embedding is a homomorphism. We next show that the embedding is one-to-one. Suppose $\nu\in M(G)$ and $\hat{\nu}=0$. Then

$$(\hat{\nu}(L)\psi,\psi)=\int_G (L_x\psi,\psi)\,d\nu(x)=0,$$

for all L in G^o and ψ in $\mathfrak{H}$. Note that every continuous positive definite function on G is of the form $x\rightarrow(L_x\psi,\psi)$, for some L in G^o and ψ in $\mathfrak{H}$. Hence $\int_G g(x)\,d\nu(x)=0$, for all continuous positive definite functions g on G. Theorem 10 of [8] implies that $\nu=0$.

Since $L^1(G)$ is contained, isomorphically and isometrically in $M(G)$, we have also verified that $L^1(G)$ is embedded isomorphically in $\mathfrak{A}(G)$. We leave to the reader the verification that the involution is preserved.

If one renorms $L^1(G)$ by $\|f\|_{C^*}=\operatorname{Sup}\|T_f\|$, where T runs through all no-where trivial $*$-representations of $L^1(G)$ acting on $\mathfrak{H}$, the embedding $f\rightarrow\hat{f}$ becomes an isometry. Indeed, each such T arises from some representation T' in G^o in such a way that $T_f=\hat{f}(T')$. Hence $\|f\|_{C^*}=\operatorname{Sup}_T\|T_f\|$ $=\operatorname{Sup}\{\|\hat{f}(T')\|:T'\in G^o\}=\|\hat{f}\|$. Since $\mathfrak{A}(G)$ is complete the mapping $f\rightarrow\hat{f}$ can be extended uniquely to $C^*(G)$, the completion of $L^1(G)$ under the norm $\|\ \|_{C^*}$. Thus $C^*(G)$ is embedded isomorphically and isometrically in $\mathfrak{A}(G)$.

THEOREM 2.3. *The group G may be embedded in $\mathfrak{A}(G)$ as follows. If x is an element of G, define $\hat{x}$ in $\mathfrak{A}(G)$ by $\hat{x}(L)=L_x$ for all L in G^o. Let $\hat{G}$ denote the image of G under the map $x\rightarrow\hat{x}$. Then $\hat{G}$ is a group of unitary elements of $\mathfrak{A}(G)$. Relative to the norm topology of $\mathfrak{A}(G)$, $\hat{G}$ is a discrete group. Relative to the weak topology of $\mathfrak{A}(G)$, $\hat{G}$ is a topological group isomorphic and homeomorphic to G.*

Proof. The verification that $x \rightarrow \hat{x}$ maps G isomorphically onto $\hat{G}$ is left to the reader. The elements of $\hat{G}$ are unitary since, if $\hat{x}$ is contained in $\hat{G}$, then

$$(\hat{x} \cdot \hat{x}^*)(L) = \hat{x}(L) \cdot \hat{x}(L)^* = L_x L_x^* = I$$

where I denotes the identity operator on $\mathfrak{H}$.

We next show that $\hat{G}$ is discrete in the norm topology. Indeed we shall show that if x and y are two distinct points of G, then $\| \hat{x} - \hat{y} \| \geqq 1$. Let R denote the left regular representation of G, μ left invariant Haar measure on G, and let U denote a neighborhood of the identity e of G such that $yU \cap xU$ is empty and where $0 \neq \mu(U) < +\infty$. Such a neighborhood can be found since G is locally compact and Hausdorff. Let $\alpha = (\mu(U))^{-\frac{1}{2}}$ and define f in $L^2(G)$ by $f = \alpha \chi_{xU}$, where χ_{xU} is the characteristic function of the set xU. Then

$$\| f \|^2 = \int_G | f(\xi) |^2 \, d\mu(\xi) = \alpha^2 \int_{xU} d\mu = 1.$$

Thus $\| f \| = 1$. Let M denote an element of G^c which is unitary equivalent to R. Then

$$\begin{aligned} \| \hat{x} - \hat{y} \| &= \operatorname{Sup} \{ \| \hat{x}(L) - \hat{y}(L) \| : L \in G^c \} \\ &\geqq \| \hat{x}(M) - \hat{y}(M) \| \\ &\geqq \| R_x f - R_y f \| \\ &\geqq \Big[\int_G | R_x f(\xi) - R_y f(\xi) |^2 \, d\mu(\xi) \Big]^{\frac{1}{2}} \\ &\geqq \Big[\int_U | f(x\xi) - f(y\xi) |^2 \, d\mu(\xi) \Big]^{\frac{1}{2}} \\ &\geqq \Big[\alpha^2 \int_U d\mu(\xi) \Big]^{\frac{1}{2}} = 1. \end{aligned}$$

The fact that G and $\hat{G}$ are homeomorphic relative to the weak topology follows immediately from the following lemma.

LEMMA 2.4. *Let G be a locally compact group. Then the topology of G is equal to the weak topology associated with the set of positive definite functions on G. More precisely, the given topology of G is the same as the smallest topology such that the positive definite functions on G are continuous.*

Proof. Let U denote a neighborhood of the identity e of the group G, whose closure is compact. Choose a neighborhood V of e such that $V = V^{-1}$

and $V^2 \subset U$. Let χ_V denote the characteristic function of the set V. Define the function f_U on G by the formula

$$f_U(x) = \int_G \chi_V(x\xi)\chi_{\check{V}}(\xi)\, d\mu(\xi)$$

for all x in G. The reader may verify that f_U is a positive definite function on G, which vanishes off of U.

Let $\mathcal{P}$ denote the collection of all left translates of functions of the form f_U, where U is a neighborhood of the identity, with compact closure. According to Proposition 3, page 23 of [8], every element of $\mathcal{P}$ is a linear combination of positive definite functions on G. Thus it is sufficient to show that the given topology of G is the weak topology associated with the collection $\mathcal{P}$.

By choosing the neighborhood U sufficiently small, it is easy to verify that the collection $\mathcal{P}$ distinguishes points. Further the functions of $\mathcal{P}$ all vanish off compact sets, and not all vanish at any one point of G. By Theorem 5G of [10], the given topology of G is the same as the weak topology determined by the collection of functions $\mathcal{P}$.

3. The identification of $\mathcal{A}(G)$, as a Banach space, with the second conjugate space of $C^*(G)$. For notational convenience, let $\mathcal{A}$ denote the group C^*-algebra, $C^*(G)$. Then the conjugate of $\mathcal{A}$, denoted $\mathcal{A}^*$, is the normed linear space of all bounded complex linear functionals on $\mathcal{A}$.

We define a map of $\mathcal{A}^{**}$ into $\mathcal{A}(G)$ as follows. If F is contained in $\mathcal{A}^{**}$, let $\hat{F}$ denote the element of $\mathcal{A}(G)$ defined by the requirement that, for each L in G^c and ψ, ϕ in $\mathfrak{H}$,

$$(\hat{F}(L)\psi, \phi) = F(f_{L,\psi,\phi})$$

where $f_{L,\psi,\phi}$ is that element of $\mathcal{A}^*$ defined by

$$f_{L,\psi,\phi}(x) = \int_G x(\xi)\,(L_\xi\psi, \phi)\, d\mu(\xi)$$

for all x in $L^1(G)$. The function $f_{L,\psi,\phi}$ is defined for all points in the completion $\mathcal{A}$ of $L^1(G)$ by continuity. (Cf. Lemma 1.4 of [6].)

THEOREM 3.1. *The map $F \to \hat{F}$ is a one-to-one linear norm decreasing map of $\mathcal{A}^{**}$ onto $\mathcal{A}(G)$. In fact, $\| \hat{F} \| \leqq \| F \| \leqq 2\| \hat{F} \|$ for all F in $\mathcal{A}^{**}$. In addition the mapping is a homeomorphism relative to the weak*-topology of $\mathcal{A}^{**}$ and the weak topology of $\mathcal{A}(G)$.*

Proof. We first verify that if F is an element of $\mathcal{A}^{**}$, then $\hat{F}$ is contained in $\mathcal{Q}(G)$. We leave to the reader the simple verification that for each fixed L in G^o and F in $\mathcal{A}^{**}$, $B(\psi,\phi)=F(f_{L,\psi,\phi})$ is a bounded bilinear functional on $\mathfrak{H}$. In fact, $|B(\psi,\phi)|\leqq\|F\|\cdot\|\psi\|\cdot\|\phi\|$ for all ψ, ϕ in $\mathfrak{H}$. Thus for each L in G^o and F in $\mathcal{A}^{**}$, $\hat{F}(L)$ is a linear operator on $\mathfrak{H}$ and $\|\hat{F}(L)\|\leqq\|F\|$, for every L in G^o. Thus $\|\hat{F}\|\leqq\|F\|$ and $\hat{F}$ satisfies condition (1) of the definition of $\mathcal{Q}(G)$. We leave to the reader the non-trivial verification that $\hat{F}$ satisfies condition (2) of Definition 1.1 and thus that $\hat{F}$ is contained in $\mathcal{Q}(G)$, as well as the trivial verification that the map $F\rightarrow\hat{F}$ preserves the linear structure.

We next proceed to show the mapping is one-to-one. Due to the correspondence between unitary representations of G and no-where trivial *-representations of $\mathcal{A}$, it follows that every positive linear form on $\mathcal{A}$ is of the form $f_{L,\psi,\psi}$, where $\psi\in\mathfrak{H}$ and $L\in G^o$. According to Theorem 1 of [9], these positive forms generate $\mathcal{A}^*$. Suppose F is contained in $\mathcal{A}^{**}$ and $\hat{F}=0$. Then $(\hat{F}(L)\psi,\psi)=F(f_{L,\psi,\psi})=0$, for all L in G^o and ψ in $\mathfrak{H}$. Thus $F=0$.

We next verify that $\|F\|\leqq 2\|\hat{F}\|$, for all F in $\mathcal{A}^{**}$. Let $\mathcal{A}_p^*$ denote the positive elements of $\mathcal{A}^*$ and $\mathcal{A}_h^*$ the hermitian elements of $\mathcal{A}^*$. For each F in $\mathcal{A}^{**}$, let

$$\|F\|_p=\operatorname{Sup}\{|F(f)|:f\in\mathcal{A}_p^*,\|f\|=1\}$$

and

$$\|F\|_h=\operatorname{Sup}\{|F(f)|:f\in\mathcal{A}_h^*,\|f\|=1\}.$$

Then clearly $\|F\|_p\leqq\|F\|_h\leqq\|F\|$.

According to Theorem 1 of [9], each element f of $\mathcal{A}_h^*$ may be expressed uniquely in the form $f=f_1-f_2$, where f_1 and f_2 are in $\mathcal{A}_p^*$. Further $\|f\|=\|f_1\|+\|f_2\|$. It is easy to verify that this fact implies $\|F\|_p=\|F\|_h$, for all F in $\mathcal{A}^{**}$.

Each f in $\mathcal{A}^*$ may be expressed in the form $f=f_1+if_2$, where f_1, f_2 are in $\mathcal{A}_h^*$. Thus

$$\begin{aligned}\|F\|&=\operatorname{Sup}\{|F(f_1)+iF(f_2)|:f_1,f_2\in\mathcal{A}_h^*,\|f_1+if_2\|\leqq 1\}\\&\leqq\operatorname{Sup}\{|F(f_1)|+|F(f_2)|:f_1,f_2\in\mathcal{A}_h^*,\|f_1\|+\|f_2\|\leqq 2\}\\&\leqq 2\operatorname{Sup}\{|F(f_1)|:f_1\in\mathcal{A}_h^*,\|f_1\|\leqq 1\}=2\|F\|_h.\end{aligned}$$

Hence

$$\begin{aligned}\|F\|\leqq 2\|F\|_p&=2\operatorname{Sup}\{|F(f)|:f\in\mathcal{A}_p^*,\|f\|=1\}\\&=2\operatorname{Sup}\{|(\hat{F}(L)\psi,\psi)|:L\in G^o,\psi\in\mathfrak{H},\|\psi\|=1\}\\&\leqq 2\operatorname{Sup}\{\|\hat{F}(L)\|:L\in G^o\}=2\|\hat{F}\|.\end{aligned}$$

We next verify that the mapping is onto. Let J be an element of $\mathfrak{A}(G)$. Define the complex functional $\mathfrak{J}$ an $\mathcal{A}_p{}^*$ by $\mathfrak{J}(f_{L,\psi,\psi}) = (J(L)\psi,\psi)$, where L is contained in G^c and ψ is an element of $\mathfrak{H}$. $\mathfrak{J}$ may be extended to all of $\mathcal{A}^*$ by linearity, since every element of $\mathcal{A}^*$ is a unique linear combination of elements of $\mathcal{A}_p{}^*$. (Cf. Theorem 1 of [9].) To show that $\mathfrak{J}$ is linear on $\mathcal{A}^*$, it is sufficient to verify that $\mathfrak{J}$ is linear on $\mathcal{A}_p{}^*$. We leave this verification to the reader. Further $\mathfrak{J}$ is bounded since, for all f in $\mathcal{A}_p{}^*$, $\|f\| \leqq 1$, $|\mathfrak{J}(f)| \leqq \|J\|$. Thus $\|\mathfrak{J}\| \leqq 2\|\mathfrak{J}\|_p \leqq 2\|J\|$. Hence $\mathfrak{J}$ is an element of $\mathcal{A}^{**}$ and $J = \hat{\mathfrak{J}}$.

The map $F \to \hat{F}$ is a homeomorphism relative to the weak *-topology of $\mathcal{A}^{**}$ and the weak topology of $\mathfrak{A}(G)$. Indeed since every element of $\mathcal{A}^*$ is a finite linear combination of elements of $\mathcal{A}_p{}^*$ (cf. Theorem 1 of [9]) a net F_λ in $\mathcal{A}^{**}$ converges to F in $\mathcal{A}^{**}$ in the weak*-topology if and only if $F_\lambda(f_{L,\psi,\psi}) = (\hat{F}_\lambda(L)\psi,\psi)$ converges to $F(f_{L,\psi,\psi}) = (\hat{F}(L)\psi,\psi)$, for every L in G^c and ψ in $\mathfrak{H}$. Thus the net F_λ converges to F in the weak *-topology of $\mathcal{A}^{**}$ if and only if the net $\hat{F}_\lambda$ converges to $\hat{F}$ in the weak topology of $\mathfrak{A}(G)$.

COROLLARY 3.2. *Each of the group algebras* $L^1(G)$, $C^*(G)$ *and* $M(G)$, *is weakly dense in* $\mathfrak{A}(G)$.

Proof. In the usual way $C^*(G) = \mathcal{A}$ may be embedded in $\mathcal{A}^{**}$ by a canonical linear isometry, $x \to x'$. Let $\mathcal{A}'$ denote the image of $\mathcal{A}$ under this embedding. Then (cf. Theorem 4, page 19 of [1]) $\mathcal{A}'$ is weak* dense in $\mathcal{A}^{**}$. By Theorem 2.2, $\mathcal{A}$ may also be embedded isometrically in $\mathfrak{A}(G)$. Let $\hat{\mathcal{A}}$ denote the image of $\mathcal{A}$ in $\mathfrak{A}(G)$. We leave to the reader the simple verification that these embeddings commute. Thus $\hat{\mathcal{A}}' = \hat{\mathcal{A}}$. Hence $C^*(G)$ is weakly dense in $\mathfrak{A}(G)$. Since $L^1(G)$ is dense in $C^*(G)$, relative to the norm of $C^*(G)$, we have that $L^1(G)$ is weakly dense in $C^*(G)$ and hence in $\mathfrak{A}(G)$. Since $L^1(G) \subset M(G)$, $M(G)$ is also weakly dense in $\mathfrak{A}(G)$.

Remark. The approximate identity of $L^1(G)$, converges weakly to the identity of $\mathfrak{A}(G)$.

Remark. We have observed many natural embeddings and identifications among the objects so far considered in this paper. Thus $L^1(G)$ is embedded in $C^*(G)$ and also in $M(G)$. G is embedded in $M(G)$ as well as in $\mathfrak{A}(G)$. $L^1(G)$, $C^*(G)$ and $M(G)$ are all embedded in $\mathfrak{A}(G)$. $C^*(G)$ is embedded in $\mathcal{A}^{**}$. The reader may verify that all these natural embeddings commute. For convenience we shall therefore make certain identifications, whenever there is no danger of confusion. Thus, as in the statement of the previous corollary, we shall consider G, $L^1(G)$, $C^*(G)$ and $M(G)$ to be subsets of $\mathfrak{A}(G)$ and also, by the corespondence between $\mathscr{A}^{**}$ and $\mathfrak{A}(G)$, as subsets of $\mathscr{A}^{**}$. This identification is summarized in the diagram below.

$$\begin{tikzcd}
 & \mathfrak{A}(G) & \\
M(G) \arrow[ur, hook] & & C^*(G) \arrow[ul, hook'] \\
G \arrow[u, hook] & L^1(G) \arrow[ul, hook'] \arrow[ur, hook] &
\end{tikzcd}$$

Remark. Since the natural embeddings of $C^*(G)$ into $\mathfrak{A}(G)$ and $\mathscr{A}^{**}$ are both isometries, it follows that the restriction of the correspondence $F \to \hat{F}$ of Theorem 3.1, to $C^*(G)$, is an isometry. More precisely, if $C^*(G)$ is considered to be a subset of $\mathscr{A}^{**}$ and if $F \to \hat{F}$ denotes the natural mapping of $\mathscr{A}^{**}$ onto $\mathfrak{A}(G)$, then $\| F \| = \| \hat{F} \|$, for all F in $C^*(G)$.

THEOREM 3.3. *The unit ball $\mathfrak{A}_1(G)$, is weakly compact.*

Proof. If J_λ is a net in $\mathfrak{A}_1(G)$, converging weakly to J in $\mathfrak{A}(G)$, then, for each L in G^c, $J_\lambda(L)$ is a net converging weakly to $J(L)$. Since $J_\lambda(L)$ is contained in the unit ball, $\mathscr{L}_1(\mathfrak{H})$, of the algebra of operators on $\mathfrak{H}$ and since $\mathscr{L}_1(\mathfrak{H})$ is weakly closed, it follows that $J(L)$ is contained in $\mathscr{L}_1(\mathfrak{H})$, for all L in G^c. Thus J is contained in $\mathfrak{A}_1(G)$. Hence $\mathfrak{A}_1(G)$ is weakly closed.

Let $\mathscr{A}_2^{**}$ denote the ball of $\mathscr{A}^{**}$ of radius 2. $\mathscr{A}_2^{**}$ is weak* compact in $\mathscr{A}^{**}$. (Cf. Corollary 4, page 41 of [1].) By Theorem 3.1, $\hat{\mathscr{A}}_2^{**}$ is weakly compact, and contains the weakly closed set $\mathfrak{A}_1(G)$. Thus $\mathfrak{A}_1(G)$ is weakly compact.

4. A remark on the choice of axioms for $\mathfrak{A}(G)$.

The reader should note that the choice of the two axioms used to define the elments of $\mathfrak{A}(G)$ was strongly motivated by the work of Tadao Tannaka. In his paper, [14], Tannaka considered certain mappings of the set of finite dimensional representations of a compact group into matrices, satisfying a certain list of axioms. On the basis of his list it would appear natural to require certain properties in addition to those specified in Definition 1.1. The following theorem states that these are already implied by the two axioms already specified.

THEOREM 4.1. *Each option J in $\mathfrak{A}(G)$ satisfies the following two properties.*

(1) *If L is an element of G^c and U is a unitary operator on $\mathfrak{H}$, then $J(U^{-1}LU) = U^{-1}J(L)U$.*

(2) *If L is an element of G^c, then $J(L)$ is contained in the von Neumann algebra $\mathfrak{A}(L)$, generated by the range of L.*

Proof. (1) For each ψ in $\mathcal{H}$, we have that the elements $f_{L,U\psi,U\psi}$ and $f_{U^{-1}LU,\psi,\psi}$ of $\mathcal{A}^*$ are identical. Let $\mathfrak{J}$ denote the element of $\mathcal{A}^{**}$ for which $\hat{\mathfrak{J}} = J$. Then $\mathfrak{J}(f_{L,U\psi,U\psi}) = \mathfrak{J}(f_{U^{-1}LU,\psi,\psi})$ and thus

$$(J(L)U\psi, U\psi) = (J(U^{-1}LU)\psi, \psi)$$

for all ψ in $\mathcal{H}$. Hence $U^{-1}J(L)U = J(U^{-1}LU)$.

(2) Choose an element L in G^c. Then the map $L' : J \rightarrow J(L)$ maps $\mathfrak{A}(G)$ into $\mathcal{L}(\mathcal{H})$. By the definition of the weak topology of $\mathfrak{A}(G)$, L' is continuous relative to the weak topology of $\mathfrak{A}(G)$ and the weak operator topology of $\mathcal{L}(\mathcal{H})$. Since $C^*(G)$ is weakly dense in $\mathfrak{A}(G)$, we conclude that $L'(C^*(G))$ is weakly dense in $L'(\mathfrak{A}(G))$. Hence $\mathfrak{A}(L)$, which is the weak closure of $L'(C^*(G))$, contains $L'(\mathfrak{A}(G))$.

5. Other topologies for $\mathfrak{A}(G)$.

Remark on Terminology. A number of topologies have been considered on the collection $\mathcal{L}(\mathcal{H})$ of bounded linear operators on a Hilbert space $\mathcal{H}$. In addition to the uniform, the strong and the weak operator topologies, John von Neumann [12] considered the locally convex topology determined by the family of semi-norms on $\mathcal{L}(\mathcal{H})$ of the form

$$T \rightarrow [\sum_{i=1}^{\infty} \| T\psi_i \|^2]^{\frac{1}{2}}$$

where $\{\psi_i\}$ denotes a sequence in $\mathcal{H}$ such that $\sum_{i=1}^{\infty} \| \psi_i \|^2 < +\infty$. Von Neumann called this the "strongest" topology on $\mathcal{L}(\mathcal{H})$. This terminology does violence to the English language, in that this topology is strictly weaker than the uniform topology, whenever $\mathcal{H}$ is infinite dimensional. Perhaps because of this fact, Jacques Dixmier chose to refer to this topology by the term "ultra-forte" in the French language (cf. [2]). In the paper [2], Jacques Dixmier considers another locally convex topology on $\mathcal{L}(\mathcal{H})$, determined by the family of semi-norms on $\mathcal{L}(\mathcal{H})$ of the form

$$T \rightarrow | \sum_{i=1}^{\infty} (T\psi_i, \phi_i) |$$

where $\{\psi_i\}$ and $\{\phi_i\}$ are sequences in $\mathcal{H}$, such that $\sum_{i=1}^{\infty} \| \psi_i \|^2 < +\infty$ and $\sum_{i=1}^{\infty} \| \phi_i \|^2 < +\infty$. In a latter paper, [3], Professor Dixmier, by reason of symmetry, introduces the term "ultra-faible" to refer to this topology. How should the terms "ultra-forte" and ultra-faible" be translated into English? If one requires the translation to be invertible, "ultra-forte" should revert back to "strongest" and, by symmetry, "ultra-faible" would become "weakest." However this terminology is completely unsatisfactory, for not only is the "ultra-faible" topology not the weakest topology, it is in general strictly stronger than the weak topology. For this same reason, the use of the English term "ultra-weak" for the French term "ultra-faible" is unsatisfactory.

The author wishes, in this section, to introduce analogues of each of these topologies, for $\mathcal{A}(G)$. Since we do not wish the perpetuate this confusing terminology in a new context, we shall offer a new term to replace the term "strongest operator topology," as well as an English term to be used in translating the French term "ultra-faible."

Convention. Henceforth the strongest topology in $\mathcal{L}(\mathcal{H})$ (in French "ultra-forte") will be referred to as the *σ-strong topology.* The French term "ultra-faible" will be translated "σ-weak."

Definition 5.1. Consider the collection of all semi-norms on $\mathcal{A}(G)$ of the form $J \to \| J(L)\psi \|$, where ψ is a vector in $\mathcal{H}$, and L is an element of G^c. This collection of semi-norms defines on $\mathcal{A}(G)$ a locally convex Hausdorff topology which we call the *strong topology* of $\mathcal{A}(G)$. A neighborhood base of the origin is given by the collection of neighborhoods of the form

$$\mathcal{N}(0;L^1,\cdots,L^n,\psi_1,\cdots,\psi_n) = \{J : J \in \mathcal{A}(G) \text{ and } \| J(L^i)\psi_i \| \leqq 1 \text{ for } i = 1,\cdots,n\}$$

where $L^i \in G^c$ and $\psi_i \in \mathcal{H}$, $i = 1,\cdots,n$.

This topology is characterized as the smallest topology of $\mathcal{A}(G)$ such that, for each L in G^c, the map $J \to J(L)$, of $\mathcal{A}(G)$ into $\mathcal{L}(\mathcal{H})$ is continuous relative to the strong operator topology.

Definition 5.2. Let $\{\psi_i\}$ denote a sequence of elements of $\mathcal{H}$ such that $\sum_{i=1}^{\infty} \| \psi_i \|^2 < +\infty$ and let L^i denote a sequence of elements of G^c. Then for each J in $\mathcal{A}(G)$, we have

$$\sum_{i=1}^{\infty} \| J(L^i)\psi_i \|^2 \leqq \| J \|^2 \sum_{i=1}^{\infty} \| \psi_i \|^2 < +\infty.$$

Thus the function

$$J \to [\sum_{i=1}^{\infty} \| J(L^i)\psi_i \|^2]^{\frac{1}{2}}$$

is a semi-norm on $\mathcal{A}(G)$. The set of all such semi-norms on $\mathcal{A}(G)$ defines a Hausdorff locally convex topology which we call the *σ-strong topology* of $\mathcal{A}(G)$. A neighborhood base of the origin is given by the collection of neighborhoods of the form

$$\mathcal{N}(0;\{L^i\},\{\psi_i\}) = \{J : J \in \mathcal{A}(G) \text{ and } (\sum_{i=1}^{\infty} \| J(L^i)\psi_i \|^2) \leqq 1\}$$

where $\{L^i\}$ is a sequence in G^c and $\{\psi_i\}$ is a sequence in $\mathcal{H}$ such that $\sum_{i=1}^{\infty} \| \psi_i \|^2 < +\infty$.

Definition 5.3. Let $\{\psi_i\}$ and $\{\phi_i\}$ denote two sequences of elements of $\mathcal{H}$ such that $\sum_{i=1}^{\infty} \| \psi_i \|^2 < +\infty$ and $\sum_{i=1}^{\infty} \| \phi_i \|^2 < +\infty$. Let $\{L^i\}$ denote a sequence in G^c. Then for each J in $\mathcal{A}(G)$, the function

$$J \to | \sum_{i=1}^{\infty} (J(L^i)\psi_i, \phi_i) |$$

is a semi-norm on $\mathcal{A}(G)$. The set of all such semi-norms defines a Hausdorff locally convex topology on $\mathcal{A}(G)$, which we call the *σ-weak topology* of $\mathcal{A}(G)$.

The collection of semi-norms of the form

$$J \to | \sum_{i=1}^{\infty} (J(L^i)\psi_i, \psi_i) |$$

where $\{\psi_i\}$ is in $\mathcal{H}$, and $\sum_{i=1}^{\infty} \| \psi_i \|^2 < +\infty$, suffices to define this topology. A neighborhood base of the origin is given by finite intersections of neighborhoods of the form

$$\mathcal{N}(0;\{L^i\},\{\psi_i\}) = \{J : J \in \mathcal{A}(G) \text{ and } | \sum_{i=1}^{\infty} (J(L^i)\psi_i, \psi_i) | \leqq 1\}$$

where $\{L^i\}$ is a sequence in G^c and $\{\psi_i\}$ is a sequence in $\mathcal{H}$ such that $\sum_{i=1}^{\infty} \| \psi_i \|^2 < +\infty$.

Let the symbol $\supset$ designate the term "contains." Then we have the following diagram of the topologies of $\mathcal{A}(G)$.

$$\begin{array}{ccccc} \text{uniform topology} & \supset & \sigma\text{-strong topology} & \supset & \text{strong topology} \\ & & \cup & & \cup \\ & & \sigma\text{-weak topology} & \supset & \text{weak topology}. \end{array}$$

The techniques developed in the study of the corresponding topologies for $\mathcal{L}(\mathcal{H})$ (cf. Section 3, Chapter 1 of [4]) may be applied in this context to obtain the following unsurprising results. The linear operations, as well as the maps of $\mathcal{A}(G)$ into $\mathcal{A}(G)$ of the form $J \to JK$ and $J \to KJ$ where K is an element of $\mathcal{A}(G)$, are continuous in any one of the five topologies. The involution $J \to J^*$ is continuous in the weak and the σ-weak topologies. The map $(J, K) \to JK$ of $\mathcal{A}_1(G) \times \mathcal{A}(G)$ into $\mathcal{A}(G)$ is continuous in the strong and σ-strong topologies.

THEOREM 5.4. *The weak and the σ-weak topologies of $\mathcal{A}(G)$ coincide on bounded sets.*

Proof. Let $\mathcal{A}_r(G) = \{J : J \in \mathcal{A}(G) \text{ and } \| J \| \leqq r\}$ where r is a positive number. We will show that the weak and σ-weak topologies coincide on $\mathcal{A}_r(G)$.

Suppose $\mathfrak{M}$ is a set of the form

$$\mathfrak{M} = \{J : J \in \mathcal{A}(G) \text{ and } |\sum_{i=1}^{\infty} (J(L^i)\psi_i, \psi_i)| \leqq 1\}$$

where $\{L^i\}$ is a sequence in G^o and $\{\psi_i\}$ is a sequence in $\mathfrak{H}$ such that $\sum_{i=1}^{\infty} \| \psi_i \|^2 < +\infty$. Choose an integer n such that $\sum_{i=n+1}^{\infty} \| \psi_i \|^2 \leqq 1/2r$. Let

$$\mathfrak{M}' = \{J : J \in \mathcal{A}(G) \text{ and } |(J(L^i)(\sqrt{2n}\psi_i), (\sqrt{2n}\psi_i))| \leqq 1, i = 1, \cdot \cdot, n\}$$

Then if $J \in \mathfrak{M}' \cap \mathcal{A}_r(G)$, we have

$$\begin{aligned} |\sum_{i=1}^{\infty} (J(L^i)\psi_i, \psi_i)| &\leqq \sum_{i=1}^{n} |(J(L^i)\psi_i, \psi_i)| + \sum_{i=n+1}^{\infty} |(J(L^i)\psi_i, \psi_i)| \\ &\leqq n(1/2n) + \| J \| \cdot \sum_{i=n+1}^{\infty} \| \psi_i \|^2 \\ &\leqq 1/2 + r(1/2r) = 1. \end{aligned}$$

Hence $\mathfrak{M}' \cap \mathcal{A}_r(G) \subset \mathfrak{M} \cap \mathcal{A}_r(G)$.

Every fundamental σ-weak neighborhood $\mathfrak{N}$ of the origin is of the form $\mathfrak{N} = \bigcap_{j=1}^{m} \mathfrak{M}_j$ where $\mathfrak{M}_j$ is a set of the form $\mathfrak{M}$ described above and m is a positive integer. By the above argument, corresponding to each $\mathfrak{M}_j$, there is a weak neighborhood $\mathfrak{M}_j'$ of the origin such that

$$\mathfrak{M}_j' \cap \mathcal{A}_r(G) \subset \mathfrak{M}_j \cap \mathcal{A}_r(G).$$

Thus $\mathfrak{N}' = \bigcap_{j=1}^{m} \mathfrak{M}_j'$ is a weak neighborhood of the origin and

$$\mathfrak{N}' \cap \mathcal{A}_r(G) = (\bigcap_{j=1}^{m} \mathfrak{M}_j') \cap \mathcal{A}_r(G) \subset (\bigcap_{j=1}^{m} \mathfrak{M}_j) \cap \mathcal{A}_r(G) = \mathfrak{N} \cap \mathcal{A}_r(G).$$

THEOREM 5.5. *The strong and the σ-strong topologies of $\mathcal{A}(G)$ coincide on bounded sets.*

The proof of this theorem is completely analogous to that of the previous theorem.

THEOREM 5.6. *The restrictions of each of the topologies weak, σ-weak, strong, σ-strong, to the set $\mathcal{U}$ of unitary elements of $\mathcal{A}(G)$, are identical.*

Proof. On the basis of the previous two theorems, it is sufficient to show that the weak and the strong topologies coincide on $\mathcal{U}$.

Suppose $\{J_\lambda\}$ is a net in $\mathcal{U}$ which converges weakly to J in $\mathcal{U}$. Since the involution is weakly continuous in $\mathcal{A}(G)$ we have that J_λ^* converges to J^*. By the weak continuity of multiplication by a fixed element we have that $\{J_\lambda^* J\}$ converges weakly to $J^*J = I$ and also that $\{J^*J_\lambda\}$ converges weakly to $J^*J = I$. Hence for each ψ in $\mathcal{H}$ and each L in G^c,

$$\begin{aligned}\|(J_\lambda - J)(L)\psi\|^2 &= ([J_\lambda(L) - J(L)]^*[J_\lambda(L) - J(L)]\psi, \psi)\\ &\leqq |((I - J^*(L)J_\lambda(L))\psi, \psi)|\\ &\quad + |((I - J_\lambda^*(L)J(L)\psi, \psi)|.\end{aligned}$$

Hence the fact that J^*J_λ and J_λ^*J both converge weakly to I, implies that J_λ converges that J converges strongly to J.

COROLLARY 5.7. *$\mathcal{U}$ is a Hausdorff topological group relative to the weak (equivalently the σ-weak, strong, or σ-strong) topologies.*

Proof. The inverse operation is continuous since the involution is weakly continuous. Multiplication is continuous since multiplication on bounded sets is strongly continuous.

COROLLARY 5.8. *The restriction of any of the four topologies, weak, σ-weak, strong, σ-strong, to G, is identical with the given topology of G.*

Proof. This follows immediately from Theorem 5.6 and Theorem 2.3.

6. The representation of $\mathcal{A}(G)$ as a Von Neumann algebra. Grothendieck [9] and Takeda [13] have developed a way of representing the Banach space $\mathcal{A}^{**}$ as linear operators on a Hilbert space. On the basis of this representation, $\mathcal{A}^{**}$ is given a $*$-algebra structure such that $\mathcal{A}^{**}$ is a von Neumann algebra. In our case, however, $\mathcal{A}(G)$ is given as a C^*-algebra. The technique of Grothendieck and Takeda is easily adapted to show that $\mathcal{A}(G)$ may be represented isomorphically and isometrically as a weakly closed

*-algebra of operators, on an (in general non-separable) Hilbert space H. The isomorphism is a σ-strong homeomorphism.

Definition of the Representation ϕ *of* $\mathfrak{A}(G)$. Let $H = \sum_{L \in G^o} \oplus \mathfrak{H}_L$ denote the Hilbert space direct sum of α copies of $\mathfrak{H}$, where α denotes the cardinal of G^o and $\mathfrak{H}_L = \mathfrak{H}$, for every L in G^o. Thus a vector ψ in H is a vector valued function on G^o, $L \to \psi_L$ such that ψ_L is contained in $\mathfrak{H}$ for each L in G^o, where only a countable number of the ψ_L are non-zero and $\sum_{L \in G^o} \| \psi_L \|^2 < +\infty$. For convenience in examining the components of ψ, we shall usually denote ψ by $\{\psi_L\}$.

We define the map ϕ of $\mathfrak{A}(G)$ into operators acting on H, by the requirement $\phi(J)\{\psi_L\} = \{J(L)\psi_L\}$ for all J in $\mathfrak{A}(G)$ and $\{\psi_L\}$ in H. It is easy to verify that, for each J, $\phi(J)$ is a bounded linear operator on H, and in fact $\| \phi(J) \| \leqq \| J \|$.

THEOREM 6.1. *$\phi(\mathfrak{A}(G))$ is a von Neumann algebra. ϕ is a *-algebra isomorphism and isometry of $\mathfrak{A}(G)$ onto $\phi(\mathfrak{A}(G))$. The extent to which the topologies correspond is given in the following list.*

1) *ϕ is a homeomorphism relative to the σ-strong topology of $\mathfrak{A}(G)$ and the σ-strong operator topology of $\phi(\mathfrak{A}(G))$.*

2) *ϕ is a homeomorphism relative to the σ-weak topology of $\mathfrak{A}(G)$ and the σ-weak operator topology of $\phi(\mathfrak{A}(G))$.*

3) *ϕ is an open mapping, relative to the strong topology of $\mathfrak{A}(G)$ and the strong operator topology of $\phi(\mathfrak{A}(G))$.*

4) *ϕ is an open mapping, relative to the weak topology of $\mathfrak{A}(G)$ and the weak operator topology of $\phi(\mathfrak{A}(G))$.*

Proof. We leave to the reader the verification that ϕ is a *-algebra isomorphism. ϕ is then an isometry, since every *-algebra isomorphism of a C^*-algebra into a C^*-algebra preserves norm. (Cf. page 311 of [11].)

We next show that ϕ is continuous relative to the σ-strong topology of $\mathfrak{A}(G)$ and the σ-strong operator topology of $\mathcal{L}(H)$. Let $\mathfrak{N}$ denote a fundamental σ-strong neighborhood of the origin in $\mathcal{L}(H)$. Then there exists a sequence $\{\psi_j\}$ in H such that

$$\sum_{i=1}^{\infty} \| \psi_j \|^2 < +\infty \text{ and } \mathfrak{N} = \{T : T \in \mathcal{L}(H) \text{ and } \sum_{i=1}^{\infty} \| T\psi_j \|^2 \leqq 1\}.$$

For each L in G^o, let $P(L)$ denote the projection of H onto its L-th coordinate. Thus for each integer j, there are at most a countable number of points $\{L^{ij}\}$ in G^o such that $P(L^{ij})\psi_j \neq 0$. Rearrange the double sequence of pairs $\{(L^{ij}, P(L^{ij})\psi_j)\}$ into a single sequence $\{(L^k, \phi_k)\}$. Note that

$$\sum_{k=1}^{\infty} \| \phi_k \|^2 = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \| P(L^{ij})\psi_j \|^2 = \sum_{j=1}^{\infty} \| \psi_j \|^2 < +\infty.$$

Let N denote the σ-strong neighborhood of the origin of $\mathfrak{A}(G)$ defined by

$$N = \{J : J \in \mathfrak{A}(G) \text{ and } \sum_{k=1}^{\infty} \| J(L^k)\phi_k \|^2 \leqq 1\}.$$

The reader may then verify that $\phi(N) \subset \mathfrak{N}$.

A completely similar technique may be used to show that ϕ is continuous relative to the σ-weak topology of $\mathfrak{A}(G)$ and the σ-weak operator topology of $\mathcal{L}(H)$.

We next show that $\phi(\mathfrak{A}(G))$ is a von Neumann algebra. Theorem 3.3 and Theorem 5.4 imply that the unit ball $\mathfrak{A}_1(G)$ is σ-weak compact. Since ϕ is an isometry and is σ-weak continuous, the unit ball of $\phi(\mathfrak{A}(G))$ is σ-weak compact and hence σ-weak closed. By Corollary 2, page 44 of [4], $\phi(\mathfrak{A}(G))$ is a von Neumann algebra.

We next show that ϕ is an open mapping of $\mathfrak{A}(G)$ onto $\phi(\mathfrak{A}(G))$ relative to the σ-weak topology of $\mathfrak{A}(G)$ and the σ-weak operator topology of $\phi(\mathfrak{A}(G))$. Let $\mathfrak{N}$ denote a fundamental σ-strong neighborhood of the origin of $\mathfrak{A}(G)$. Then $\mathfrak{N}$ is of the form

$$\mathfrak{N} = \{J : J \in \mathfrak{A}(G) \text{ and } \sum_{i=1}^{\infty} \| J(L^i)\psi_i \|^2 \leqq 1\}$$

where $\{\psi_i\}$ is a sequence of vectors in $\mathfrak{H}$ for which $\sum_{i=1}^{\infty} \| \psi_i \|^2 < +\infty$ and $\{L^i\}$ is a sequence of elements of G^o. Let $\bar{\psi}_i$ denote the element of H whose L^i-th-component is ψ_i, and for which every other component is zero. Let N denote the σ-strong neighborhood of the origin of $\phi(\mathfrak{A}(G))$ defined by

$$N = \{T : T \in \phi(\mathfrak{A}(G)) \text{ and } \sum_{i=1}^{\infty} \| T\bar{\psi}_i \|^2 \leqq 1\}.$$

The reader may then verify that $N \subset \phi(\mathfrak{N})$.

By a completely similar technique, one may verify that ϕ is an open mapping of $\mathfrak{A}(G)$ onto $\phi(\mathfrak{A}(G))$, relative to the strong (respectively weak, σ-weak) topology of $\mathfrak{A}(G)$ and the strong (respectively weak, σ-weak) operator topology of $\phi(\mathfrak{A}(G))$.

COROLLARY 6.2. *Let ϕ_1 denote the restriction of ϕ to the unit ball $\mathcal{A}_1(G)$ of $\mathcal{A}(G)$. Then ϕ_1 is a homeomorphism relative to the strong (respectively weak) topology of $\mathcal{A}_1(G)$ and the strong (respectively weak) operator topology of $\phi(\mathcal{A}_1(G))$.*

COROLLARY 6.3. *The classical group algebras, $L^1(G)$, $C^*(G)$ and $M(G)$ are each dense in $\mathcal{A}(G)$, relative to any one of the topologies, weak, σ-weak, strong, σ-strong, of $\mathcal{A}(G)$.*

Proof. It is sufficient to show that these classical group algebras are dense in the strongest of these topologies, namely the σ-strong topology. By Corollary 3.2, each of these group algebras is weakly dense in $\mathcal{A}(G)$. Since ϕ is an open map relative to the weak topology of $\mathcal{A}(G)$ and the weak operator topology of $\phi(\mathcal{A}(G))$, $\phi(L^1(G))$, $\phi(C^*(G))$ and $\phi(M(G))$ are each weakly dense in $\phi(\mathcal{A}(G))$. Since $\phi(\mathcal{A}(G))$ is a von Neumann algebra, von Neumann's density theorem (Theorem 2, page 43 of [4]) implies that $\phi(L^1(G))$, $\phi(C^*(G))$ and $\phi(M(G))$ are each σ-strong dense in $\phi(\mathcal{A}(G))$. Since ϕ is a σ-strong homeomorphism, we conclude that $L^1(G)$, $C^*(G)$ and $M(G)$ are each σ-strong dense in $\mathcal{A}(G)$.

COROLLARY 6.4. *Every separable locally compact group G is homeomorphic and isomorphic to a group of unitary operators acting on a Hilbert space, where the unitary group is considered as a topological group relative to any one of the four topologies, weak, σ-weak, strong, or σ-strong.*

Proof. This observation is an immediate consequence of Corollary 5.8 and Corollary 6.2.

7. The group G generates the group algebra $\mathcal{A}(G)$. When G is a finite group, the group algebra is obtained simply by taking linear combinations of group elements. Thus an algebra is associated with the group by the same procedure that one associates an algebra with the range of each representation of the group. The classical group algebras for locally compact groups have no analogous property. This $L^1(G)$ and $C^*(G)$ in general do not even contain the group, and the elements of the algebra are obtained by integration over the group. The measure group algebra $M(G)$ contains the group, but there is no natural sense in which G generates $M(G)$ as a topological algebra. In this section we shall show that G generates $\mathcal{A}(G)$ in the sense that the smallest weakly closed *-subalgebra of $\mathcal{A}(G)$, containing G, is $\mathcal{A}(G)$ itself. Thus G generates $\mathcal{A}(G)$ in the same manner in which one associates an algebra with the range of each representation of G. It is in this sense that the author argues that $\mathfrak{A}(G)$ is the natural analogue of the group algebra of a finite group.

We first note that in the isomorphic representation ϕ of $\mathfrak{A}(G)$ as a von Neumann algebra, the elements of $L^1(G)$ are related to the elements of G by the expected formula.

LEMMA 7.1. *Consider G and $L^1(G)$ as embedded in $\mathfrak{A}(G)$, according to Section 2. Let ϕ denote the representation of $\mathfrak{A}(G)$ described in Section 6. Then for all ψ in H and f in $L^1(G)$,*

$$(\phi(f)\psi,\psi) = \int_G f(x)\,(\phi(x)\psi,\psi)\,d\mu(x).$$

We leave the proof of this lemma to the reader.

THEOREM 7.2. *Let $\mathcal{A}(G)$ denote the *-algebra of elements of $\mathfrak{A}(G)$, which are complex linear combinations of elements of G. Let $\mathcal{A}_1(G)$ denote the unit ball of $\mathcal{A}(G)$. Then*

(i) *$\mathcal{A}(G)$ is dense in $\mathfrak{A}(G)$, relative to any one of the four topologies, weak, σ-weak, strong and σ-strong.*

(ii) *$\mathcal{A}_1(G)$ is dense in $\mathfrak{A}_1(G)$, relative to any one of the four topologies, weak, σ-weak, strong and σ-strong.*

Proof. Suppose T is a linear operator on H, which commutes with all the elements of $\phi(G)$. Lemma 7.1 may then be used to verify that T commutes with each element of $\phi(L^1(G))$. The double commutator theorem implies that the von Neumann algebra generated by $\mathfrak{A}(G)$, contains $\phi(L^1(G))$. Since $\phi(L^1(G))$ is weakly dense in $\phi(\mathfrak{A}(G))$, we conclude that $\phi(\mathfrak{A}(G))$ is the von Neumann algebra generated by $\phi(G)$. Thus $\phi(\mathcal{A}(G))$ is σ-strong dense in $\phi(\mathfrak{A}(G))$ and $\phi(\mathcal{A}_1(G))$ is σ-strong dense in $\phi(\mathfrak{A}_1(G))$. Since ϕ is a σ-strong homeomorphism, we have that $\mathcal{A}(G)$ is σ-strong dense in $\mathfrak{A}(G)$ and $\mathcal{A}_1(G)$ is σ-strong dense in $\mathfrak{A}_1(G)$. The theorem now follows since the σ-strong topology is the strongest of the four topologies mentioned.

COROLLARY 7.3. *$\mathfrak{A}(G)$ is commutative if and only if G is commutative.*

Notice that the corresponding fact for each of the classical group algebras follows immediately from this corollary, since each of the classical group algebras is dense in $\mathfrak{A}(G)$.

THEOREM 7.4. *The group G forms a linearly independent set in the linear topological space $\mathfrak{A}(G)$.*

Proof. Let $x_1, \cdots, x_n$ denote any finite set of distinct elements of G and suppose $\{\alpha_i\}$ is a set of complex scalars such that $\sum_{i=1}^{n} \alpha_i x_i = 0$. Then,

$$\left(\sum_{i=1}^{n} \alpha_i x_i\right)(L) = \sum_{i=1}^{n} \alpha_i L_{x_i} = 0$$

for all L in G^o. Taking L to be a representation equivalent to the regular representation we conclude that

$$\sum_{i=1}^{n} \alpha_i f(x_i^{-1}y) = 0$$

for all f in $L^2(G)$ and μ- almost all y in G, where μ denotes Haar measure on G.

Let U denote a compact neighborhood of the identity element of G, for which $x_i U \cap x_j U = \phi$, whenever $i \neq j$, where ϕ denotes the empty set. Let χ denote the characteristic function of U. Then χ is an element of $L^2(G)$ and hence $\sum_{i=1}^{n} \alpha_i \chi(x_i^{-1}y) = 0$ for μ- almost all y in G. Integrating this expression, with respect to μ, over the set $x_j U$ gives $\alpha_j \mu(U) = 0$ and hence $\alpha_j = 0$.

8. Representation theory. The main deficiency of the measure group algebra $M(G)$ is that the restriction of a $*$-representation of $M(G)$, to G, need not be strongly continuous. Thus there appears to be no reasonable analogue of the well known correspondence between the $*$-representation theory of $L^1(G)$ and $C^*(G)$, and the strongly continuous unitary representation theory of G. The purpose of this section is to produce a reasonable analogue of this situation, in the case of the algebra $\mathfrak{A}(G)$.

Definition 8.1. We say that an element J in $\mathfrak{A}(G)$ is *positive* if $(J(L)\psi, \psi) \geqq 0$ for all L in G^o and ψ in $\mathfrak{H}$. The set of positive elements of $\mathfrak{A}(G)$ will be denoted $\mathfrak{A}(G)^+$. We say $J \leqq K$, where J and K are elements of $\mathfrak{A}(G)^+$, if $(K - J) \in \mathfrak{A}(G)^+$.

The reader should have no difficulty in verifying that this partial ordering is exactly the partial ordering induced by the partial ordering among the positive elements of the von Neumann algebra $\phi(\mathfrak{A}(G))$, where ϕ is the isomorphism described in Section 6.

Definition 8.2. We define a *normal representation* of $\mathfrak{A}(G)$ to be a $*$-algebra homomorphism T into a von Neumann algebra such that, if C is a chain in $\mathfrak{A}(G)^+$, with supremum J in $\mathfrak{A}(G)^+$, then $T(J)$ is the supremum of $T(C)$.

THEOREM 8.3. *Let L be an element of G^o. Define the map L' of $\mathcal{A}(G)$ into $\mathcal{A}(L)$, the von Neumann algebra generated by the range of L, by $L'(J) = J(L)$, for all J in $\mathcal{A}(G)$. Then L' is a normal representation of $\mathcal{A}(G)$ onto $\mathcal{A}(L)$. The restriction of L' to G is L.*

If T is a normal representation of $\mathcal{A}(G)$, with representation space $\mathfrak{H}$, then the restriction of T to G is a strongly continuous unitary representation L of G, and $L' = T$.

For each I in G^c, L is irreducible (respectively primary, type I, type II, type III, finite, infinite, multiplicity free) if and only if L' is irreducible (respectively primary, type I, type II, type III, finite, infinite, multiplicity free).

For each pair of representations L and M in G^o, L is equivalent (respectively quasi-equivalent) to M if and only if L' is equivalent (respectively quasi-equivalent) to M'.

Proof. L' is easily verified to be a *-algebra homomorphism of $\mathcal{A}(G)$ into $\mathcal{A}(L)$. By the definition of the weak topology of $\mathcal{A}(G)$, the positive linear forms $J \to (J(L)\psi, \psi)$, where $\psi \in \mathfrak{H}$, are weakly continuous and hence σ-weakly continuous on $\mathcal{A}(G)$. Hence by Theorem 1, page 54 of [4], all such positive linear forms are normal. Hence the representation L' is a normal representation. By Corollary 2, page 57 of [4], $L'(\mathcal{A}(G))$ is a von Neumann algebra. Since $L'(\mathcal{A}(G))$ contains the range of L, we have $L'(\mathcal{A}(G)) = \mathcal{A}(L)$. For each x in G, $L'(x) = \hat{x}(L) = L_x$. Thus the restriction of L' to G is just L.

Suppose T is a normal representation of $\mathcal{A}(G)$, with representation space $\mathfrak{H}$. Since $\mathcal{A}(G)$ is homeomorphic to a von Neumann algebra, relative to the σ-strong topology, T is σ-strong continuous. Since the strong and the σ-strong topologies are identical on G, it follows that the restriction L of T to G is a strongly continuous unitary representation acting on the space $\mathfrak{H}$. Thus L is an element of G^o. Note that T and L' are both σ-strong continuous *-representations of $\mathcal{A}(G)$, which have identical restrictions to G. Theorem 7.2 implies that $T = L'$.

The third paragraph of the theorem is an immediate consequence of the fact that, for each L in G^o, the range of L' is $\mathcal{A}(L)$.

Suppose L and M are in G^o and $L \cong M$. Then there exists a unitary operator U on $\mathfrak{H}$ such that $L = UMU^{-1}$. Theorem 4.1 implies that, for each J in $\mathcal{A}(G)$,

$$L'(J) = J(L) = J(UMU^{-1}) = UJ(M)U^{-1} = UM'(J)U^{-1}.$$

Thus $L' = UM'U^{-1}$ and hence $L' \cong M'$. The corresponding result for quasi-equivalence follows from Lemma 4 of [5].

Remark. Since $L^1(G)$ and $C^*(G)$ are contained in $\mathfrak{A}(G)$, the *-representation theory of $L^1(G)$ and $C^*(G)$ can be obtained by restricting normal representations of $\mathfrak{A}(G)$ to these subalgebras. It is not difficult to show that all no-where trivial, separable *-representations of $L^1(G)$ and $C^*(G)$ are obtained in this way. Thus the simple process of restriction gives rise to a four-way correspondence between the normal representation theory of $\mathfrak{A}(G)$, the no-where trivial *-representation theory of $L^1(G)$ and $C^*(G)$, and the unitary representation theory of G. These correspondences commute. Thus the classical theory of the correspondence between the representations of G and of $L^1(G)$, is canonically embedded in this new theory.

The question arises as to whether normal representations are in fact the natural representations to consider when studying the big group algebra $\mathfrak{A}(G)$. The following theorem is an attempt to answer this question in the affirmative.

THEOREM 8.4. *Let T denote a bounded separable *-representation of $\mathfrak{A}(G)$. Then the following five conditions are equivalent.*

(i) *T is a normal representation.*

(ii) *T is continuous relative to the weak operator topology and the weak topology of $\mathfrak{A}(G)$.*

(iii) *T is continuous relative to the σ-weak operator topology and the σ-weak topology of $\mathfrak{A}(G)$.*

(iv) *T is continuous relative to the strong operator topology and the strong topology of $\mathfrak{A}(G)$.*

(v) *T is continuous relative to the σ-strong operator topology and the σ-strong topology of $\mathfrak{A}(G)$.*

Proof. By Theorem 8.3, each normal representation T is of the form $T = L'$, where L is the restriction of T to G. However every map of the form $J \to J(L) = L'(J)$ satisfies the continuity conditions (ii) through (v), by the definition of the different topologies for $\mathfrak{A}(G)$.

Conversely, suppose T satisfies any one of the continuity conditions (ii) through (v). Then the restriction of T to G gives a strongly continuous unitary representation L of G. Since the representation L' satisfies conditions

(ii) through (v), Theorem 7.2 implies $T = L'$. Thus Theorem 8.3 implies that T is normal.

According to Corollary 5.7, the group $\mathcal{U}$ of unitary elements of $\mathcal{A}(G)$, is a Hausdorff topological group, relative to any one of the topologies weak, σ-weak, strong and σ-strong. The restriction of any normal representation of $\mathcal{A}(G)$ gives rise to a strongly continuous unitary representation of $\mathcal{U}$. Thus Theorem 8.3 implies that the unitary representation theory of G and the unitary representation theory of $\mathcal{U}$, are identical.

THEOREM 8.3. *Every separable, strongly continuous unitary representation L of G may be extended, uniquely, to a strongly continuous unitary representation L' of $\mathcal{U}$. The von Neumann algebra $\mathcal{A}(L)$ generated by the range of L is equal to the von Neumann algebra generated by the range of L'. This correspondence between representations of G and of $\mathcal{U}$, preserves the usual representation theoretic properties. (Cf. the last two paragraphs of Theorem* 8.3.)

9. Characterization of the unitary group $\mathcal{U}$ or $\mathcal{A}(G)$.

Definition 9.1. An option J in $\mathcal{A}(G)$ is said to *preserve tensor products* if, whenever U is a linear isometry of $\mathcal{H} \otimes \mathcal{H}$ onto $\mathcal{H}$ and whenever L and M are elements of G^c, then

$$J(U(L \otimes M)U^{-1}) = U(J(L) \otimes J(M))U^{-1}.$$

An option J in $\mathcal{A}(G)$ is said to *preserve tensor powers* if, whenever U is a linear isometry of $\mathcal{H} \otimes \mathcal{H}$ onto $\mathcal{H}$ and L is an element of G^c, then

$$J(U(L \otimes L)U^{-1}) = U(J(L) \otimes J(L))U^{-1}.$$

In the Tannaka duality theorem [14], referred to in Section 4, the mappings are required to preserve tensor products. In addition, each representation must map into an invertible matrix and each unitary representation must map into a unitary matrix. In this paper, however, we are only considering unitary representations. Under this restriction, the requirement that tensor products be preserved is equivalent to the requirement that representations map into unitary operators. Indeed we obtain the following characterization of the unitary group $\mathcal{U}$ of $\mathcal{A}(G)$.

THEOREM 9.2. *Suppose J is an invertible element of $\mathcal{A}(G)$. Then the following three conditions are equivalent.*

(i) *J is unitary.*

(ii) *J preserves tensor powers.*

(iii) *J preserves tensor products.*

Proof. We first prove that (ii) implies (i). Let J denote an element of $\mathcal{A}(G)$ which preserves tensor powers.

Suppose $\| J \| > 1$. Then there exists an L in G^c such that $\| J(L) \| = \alpha > 1$. Let β denote any positive number. Then there exists an integer m (of the form 2^n) such that $\alpha^m > \beta$. Let $M = U(\prod_m \otimes L)U^{-1}$, where $\prod_m \otimes L$ denotes the tensor product of m copies of L and U is a linear isometry of the representation space of $\prod_m \otimes L$ onto $\mathfrak{H}$. Then M is an element of G^c and $\| J(M) \| = \| J(L) \|^m = \alpha^m > \beta$. Thus the function $M \to \| J(M) \|$ is not bounded on G^c, in contradiction to the first requirement in the definition of an option (cf. Definition 1.1). Thus $\| J \| \leqq 1$.

It is easy to verify that if J is invertible and preserves tensor products, then J^{-1} preserves tensor products. By the previous paragraph, we have that $\| J^{-1} \| \leqq 1$. Thus $1 = \| J^{-1}J \| \leqq \| J^{-1} \| \cdot \| J \| \leqq 1$. Hence $\| J \| = 1$ and $\| J \| = 1$.

Suppose, for some representation L in G^c, that $J(L)$ is not an isometry. Then there exists a vector ψ in $\mathfrak{H}$ such that $\| \psi \| = 1$ and $\| J(L)\psi \| = \alpha < 1$. Since $J(L)$ is invertible, we know $\alpha \neq 0$. Then $\| (1/\alpha)J(L)\psi \| = 1$ and $\| J^{-1}(L)((1/\alpha)J(L)\psi) \| = (1/\alpha) > 1$, which implies $\| J^{-1} \| > 1$, which is a contradiction. Thus for each L in G^c, $J(L)$ is an invertible linear isometry on $\mathfrak{H}$ and thus a unitary operator on $\mathfrak{H}$. Hence J is unitary.

We next prove that (i) implies (iii). Suppose M and N are two representations in G^c, and that U is a linear isometry of $\mathfrak{H}$ onto the representation space of $M \otimes N$. Let $L = U(M \otimes N)U^{-1}$. Then L is an element of G^c. Let L' denote the unitary representation of $\mathcal{U}$ given by Theorem 8.5, i.e., let $L'(J) = J(L)$ for all J in $\mathcal{U}$. Define the map Φ of $\mathcal{U}$ into unitary operators on $\mathfrak{H}$ by $\Phi(J) = U(J(M) \otimes J(N))U^{-1}$ for all J in $\mathcal{U}$. Then Φ is easily verified to be a strongly continuous unitary representation of $\mathcal{U}$, whose restriction to G is just L. By the uniqueness portion of Theorem 8.5, we have $\Phi = L'$. Thus for each J in $\mathcal{U}$, $J(U(M \otimes N)U^{-1}) = J(L) = L'(J) = \Phi(J) = U(J(M) \otimes J(N))U^{-1}$.

UNIVERSITY OF ROCHESTER.

REFERENCES.

[1] M. M. Day, *Normed linear spaces*, Springer-Verlag, Berlin, 1958.

[2] J. Dixmier, "Les fonctionnelles linéaires sur l'ensemble des opérateurs bornés d'un espace de Hilbert," *Annals of Mathematics*, vol. 51 (1950), pp. 387-408.

[3] ———, Formes linéaires sur un anneau d'operateurs," *Bulletin de la Société Mathematique de France*, vol. 81 (1953), pp. 9-39.

[4] ———, *Les algèbres d'opérateurs dans l'espace hilbertien*, Gautheir-Villars, Paris, 1957.

[5] J. Ernest, "A decomposition theory for unitary representations of locally compact groups," *Transactions of the American Mathematical Society*, vol. 104 (1962), pp. 252-277

[6] J. M. G. Fell, "The dual spaces of C^*-algebras," *Transactions of the American Mathematical Society*, vol. 94 (1960), pp. 365-403.

[7] ———, "C^*-algebras with smooth dual," *Illinois Journal of Mathematics*, vol. 4 (1960), pp. 221-230.

[8] R. Godement, "Les fonctions de type positif et la theorie des groupes," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 1-84.

[9] A. Grothendieck, "Un résultat sur le dual d'une C^*-algèbre," *Journal Mathématique Pures et Appliquées*, vol. 36 (1957), pp. 97-108.

[10] L. H. Loomis, *An introduction to abstract harmonic analysis*, D. Van Nostrand Co., New York, 1953.

[11] M. A. Naimark, *Normed rings*, P. Noordhoff, Groningen, 1959.

[12] J. von Neumann, "On a certain topology for rings of operators," *Annals of Mathematics*, vol. 37 (1936), pp. 111-115.

[13] Z. Takeda, "Conjugate spaces of operator algebras," *Proceedings of the Japan Academy*, vol. 30 (1954), pp. 90-95.

[14] T. Tannaka, "On the duality theorem for non-commutative topological groups," *Tôhoku Mathematical Journal*, vol. 45 (1939), pp. 1-12.

BRANCHED REGULAR CURVE FAMILIES AND FINITE ASYMPTOTIC VALUES.

By David V. V. Wend.[1]

Introduction. The level curves of the real part of a function analytic in a simply connected domain form a branched regular curve family, i. e., a family of curves which is locally homeomorphic to parallel lines at every point of the domain except at most at an isolated set of branch points at each of which the family has a singularity of the multiple saddle point type. Conversely, Boothby [2, 3] has shown that any given branched regular curve family filling a simply connected domain is homeomorphic to the family of level curves of a harmonic function. Such a harmonic function will be said to be contoured by the curve family. The main purpose of this paper is to investigate some relations between the structure of a branched regular curve family and the finite asymptotic values of analytic functions the real or imaginary parts of which are contoured by the given family. In particular we obtain sufficient conditions that any analytic function with real or imaginary part contoured by a given family have a certain minimum number of distinct finite asymptotic values and sufficient conditions that there exists an analytic function with no finite asymptotic value with real or imaginary part contoured by a given family. As a result we obtain precise characterizations of the level curve families of harmonic polynomials.

In Section 1 we consider some basic definitions and properties relating to finite asymptotic values and branched regular curve families including the definition of the important *b*-curves. In Section 2 we give some sufficient conditions on a branched regular curve family *F* that any analytic function with real part contoured by *F* shall have at least a certain minimum number of finite asymptotic values. In particular we prove that a *b*-curve of *F* contains the homeomorphic image of a finite asymptotic path for any analytic function with real part contoured by *F*. In Section 3 the restrictions are put on *F* that *F* contain no *b*-curve and that at most a finite number of

---

Received October 4, 1963.

Revised May 18, 1964.

[1] The material from this paper was taken from the author's Ph. D. thesis at the University of Michigan. The author wishes to express his gratitude for the assistance and encouragement given to him in this research by Professor Wilfred Kaplan.

curves of F which have a branch point as limit point intersect any given bounded subset of the plane. Then we show the existence of an analytic function with no finite asymptotic value whose real part is contoured by F. Finally we consider the level curve families of harmonic polynomials: they are homeomorphic precisely to those branched regular curve families filling the plane which contain no b-curves and have at most a finite number of branch points.

1. Preliminary definitions and theorems.

1.1. *Finite asymptotic values.* Suppose $f(z)$ is a function defined in a simply connected domain D of the complex plane and $c\colon z = h(t)$, $0 \leqq t < 1$, is a path in D such that $h(t)$ tends to the boundary of D as t tends to 1. If $\lim f(h(t))$ as $t \rightarrow 1$ exists and is finite, then the limit is said to be a *finite asymptotic value* of $f(z)$ and c is a *finite asymptotic path* of $f(z)$ [4].

Following Iversen [1, p. 271], [6] and [11], the finite asymptotic values of an analytic function approached along two paths will be considered the same if and only if they correspond to the same boundary point of the Riemann surface of the inverse function, and in the case when the values are the same the paths will be called *equivalent* finite asymptotic paths.

1.2. *Branched regular curve families.* Next we consider some definitions and properties relating to branched regular curve families. For a detailed study of such families including definitions and proofs of properties listed below the reader is referred to the papers by W. Boothby [2, 3].

1.2.1. *Curve families.* A *curve* is the homeomorphic image of an open interval, a half open interval or a closed interval. A family F of curves will be said to *fill* a subset R of the complex plane if and only if every curve of F is in R and every point of R lies on one and only one curve of F. If U is a subset of R such that each curve c of F intersects U in a set $c \cap U$ each of whose components is a curve, then we denote by $F[U]$ the curve family filling U whose curves are the components of $c \cap U$ for all c in F. If the curve family F fills R and the curve family G fills S, then F and G will be called *homeomorphic* if there is a homeomorphism of R onto S such that the image of each curve of F is a curve of G.

1.2.2. *Regular points and branch points.* Let F be a curve family filling a set R. The family F will be said to be *regular* at a point P of R and P will be said to be a *regular point* of F if there is a neighborhood $U(P)$ of P to which P is interior relative to R and such that $F[U(P)]$ is homeomorphic to the family of segments $y = c$, $-1 < x < 1$, $-1 < c < 1$. F will be called a *regular curve family* filling R if it is regular at each point of R. Let C_0 denote the unit disk $|z| < 1$. A point B on the boundary of R will be called a *branch point* if there is a positive integer n and a neighborhood $U(B)$ of B such that $F[U(b) - B]$ is homeomorphic to the level curve family of $Re(z^n)$, $n > 1$, in C_0 minus the origin under a homeomorphism h which maps $U(B)$ onto C_0 with $h(B) = 0$. The integer n is the *multiplicity* of B, and B will be called a branch point of *order n*.

Let π denote the finite complex plane. Suppose F is regular throughout a region $R = \pi - \beta$, where β consists solely of branch points. Then F will be called a *branched regular curve family.* Henceforth F will denote a *branched regular curve family* filling π minus the set of branch points. For convenience we have dropped the phrase "minus the set of branch points" and referred simply to curve families filling a simply connected domain, and we shall in general continue to do so.

Since the results we obtain depend only on the topological structure of the curve families under consideration, there will be no loss in generality in assuming that the curve families fill π minus the set of branch points.

1.2.3. *The curves, chains and trees of F.* A curve of F is the homeomorphic image of an open interval. There are two types of curves: *regular* curves which tend to infinity in each direction and *branched* curves which have a branch point as limit in at least one direction and have at most one limiting point or *endpoint,* necessarily a branch point, in any given direction.

Since F can contain no closed curve or simple polygon of branched curves [2, p. 412], there are exactly $2n$ branched curves with a given branch point of order n as endpoint. If B is a branch point, the set of branched curves of F with B as an endpoint, together with their endpoints, forms a set called the *star* of B, denoted by $St(B)$. We shall call two branched curves c and c' of $St(B)$ *clockwise adjacent* if in the order induced on $St(B)$ by the orientation of π moving clockwise around B, c' is the unique curve which follows immediately after c. If c', c form a clockwise adjacent pair, then c, c' will be *counterclockwise adjacent.*

If the union of the curves of a collection k of branched curves together with their endpoints forms a simple curve c, we call k a *chain.* If the curve c tends to infinity in both directions, then we shall call k an *infinite chain.* If one of the regions into which c separates π contains no branched curve with a limit point on c, then k is called a *maximal chain.*

Suppose c is a fixed branched curve of F. Then the collection of all curves c' of F such that there is a finite chain containing c and c' forms, together with the endpoints of the curves of the collection, a topological *tree*. For future reference we state the following theorems.

THEOREM 1.2-1. [2, p. 421] *If T is a tree, then each curve of T is contained in exactly two maximal chains which intersect only on this curve, and every branch point of order n is contained in exactly $2n$ maximal chains whose only common point is the branch point itself.*

THEOREM 1.2-2. [2, p. 421] *A maximal chain of a tree T divides π into two domains whose complete boundary it is, and one of these domains contains no point of T.*

Let c be a branched curve. By Theorem 1.2-1 there are exactly two maximal chains m and m' containing c. If we delete the open curve c from $m \cup m'$, then $(m \cup m') - c$ consists of at most two disjoint chains, each of which we shall call a *mixed maximal chain*. An *element* of F will mean a regular curve, maximal chain or mixed maximal chain of F.

Note. When we use the terms chain, infinite chain, maximal chain, element, etc., in the point set sense, we include the endpoints of the branched curves involved.

THEOREM 1.2-3. [2, p. 430] *If K is any compact subset of π, then there are at most a finite number of distinct trees of F whict intersect K on more than one curve of the tree, and no more than a finite number of curves from any one tree can intersect K.*

We say the trees of F are *isolated* if any compact set in π intersects at most a finite number of trees of F.

1.2.4. *Chordal relations and cyclic sets.* See W. Kaplan [8, 9] for a detailed discussion of chordal systems for regular curve families. Suppose c is a regular curve or infinite chain (including endpoints). Then c separates π into two complementary domains. Denote one of these domains by $D(c)$. Then the other domain will be denoted by $D^*(c)$. Suppose b, c and d are three nonintersecting elements of F. If c separates b and d, then we write $b/c/d$. If none of b, c and d separates the other two, then b, c and d form the boundary of a simply connected domain, and we write $|b, c, d|$. A nonempty set V of disjoint elements will be said to be *cyclic* if $|b, c, d|$ for every three elements b, c and d of V. If V is a cyclic set, then the elements of V form the complete boundary of a simply connected domain which will be denoted by $\theta(V)$. If V contains only one element c, then $\theta(V) = D(c)$ for arbitrary choice of $D(c)$.

1.2.5. *The b-curves of F.* If b is a regular curve or maximal chain of F, b will be called a boundary curve or *b-curve* if there exists a regular curve or maximal chain c of F such that no curve in $\theta(b,c)$ has an endpoint on b or c and no regular curve or infinite chain separates b and c. It follows that c must also be a b-curve. We call such a pair b, c a pair of *related b-curves*. Our definition generalizes to branched regular curve families the definition of b-curve given by Cullen [5, p. 15].

1.2.6. *Cross-sections.* A curve is a *cross-section* of F if and only if it is contained in π minus the set of branch points of F and intersects each chain or regular curve of F in at most one point. In particular no cross-section can intersect a maximal chain or tree twice. A cross-section x is said to *tend properly to infinity* in a given direction on x if x tends to infinity in that direction in such a manner that the curves and trees of F meeting x tend to infinity with their intersection points on x. I.e. for each compact set K in π there exists a compact set K' containing K such that if the point of intersection of a curve c of F with x lies outside K', then c and any tree containing c lie entirely outside K. If a cross-section tends to infinity in a given direction but does not tend properly to infinity in that direction, then we say it tends *improperly* to infinity in that direction.

THEOREM 1.2-4. [3, p. 530] *If P is a regular point on a regular curve or infinite chain c of F and $D(c)$ has been chosen, then there is a cross-section of F which starts at P and tends properly to infinity in $D(c)$.*

We say a curve, chain or tree c of F is *accessible* from a point P if there is a cross-section of F containing P and a point of c. A curve, chain or tree c will be accessible from another curve, chain or tree c' of F if c is accessible from some point of c'.

1.2.7. *Parallel families.* Let x be a cross-section of F. Then we denote by $F(x)$ the set of curves of F which intersect x. The union S of the curves of $F(x)$ fills an open simply connected set, and there is a homeomorphism of S onto (1) a strip $0 \leqq y \leqq 1$, (2) a half-plane $y \geqq 0$ or (3) the xy-plane, carrying the curves of S onto the lines $y =$ constant and x into the y-axis, the case depending on whether (1) x is an arc, (2) x is a half-open curve or (3) x is an open curve, respectively [3, p. 514]. The family S will be called a *parallel family* in each case.

THEOREM 1.2-5. [3, p. 531] *The boundary of a parallel family S is a collection of maximal chains, mixed maximal chains and regular curves, and a mixed maximal chain c lies on the boundary of S if and only if $D(c)$*

can be chosen to contain S, at the same time one and only one curve d of the tree to which c belongs enters $D(c)$ from a branch point of c and the curve d is an element of S.

The following theorem follows from the fact that a parallel family is homeomorphic to a set of lines filling a simply connected region.

THEOREM 1.2-6. *The boundary of a parallel family forms a cyclic set of elements of F.*

The definition of parallel family also leads to the following theorem.

THEOREM 1.2-7. *Suppose b, c and d are curves of a parallel family $S = F(x)$, where x is the curve $z = h(t)$, $0 \leqq t \leqq 1$, $0 \leqq t < 1$ or $0 < t < 1$, and $b = F(h(t_0))$, the curve of F through $z = h(t_0)$, $c = F(h(t_1))$ and $d = F(h(t_2))$, where $t_0 < t_1 < t_2$. Then $b/c/d$ in S.*

1.2.8. *Complementary families and functions contoured by them.* Suppose F and G are branched regular curve families filling π. We say G is complementary to F if (1) the branch points of G are exactly those of F and each is of the same order, i. e. a point B is an n-th order branch point of G if and only if it is an n-th order branch point of F, and (2) every curve of G is a cross-section of F. If G is complementary to F, then F is complementary to G.

THEOREM 1.2-8. [3, p. 526] *Every branched regular curve family has at least one complementary family.*

Adapting terminology used by Morse and Jenkins [7], we use the term *conjugate net* for a pair of complementary curve families, and we say a conjugate net (F, G) *contours* a function $f(z)$, or $f(z)$ is *contoured by* (F, G), if F and G are homeomorphic to the families of level curves of the real and imaginary parts of $f(z)$ respectively under the same homeomorphism. We shall also say that F contours a real-valued function $u(x, y)$ if F is homeomorphic to the family of level curves of u.

A real-valued function $u(x, y) = u(P)$ is called an *allowable function* contoured by F if

(1) $u(P)$ is defined and continuous for all P in π;

(2) for every real number k the locus $u(P) = k$, if not empty, consists of a finite or an at most countably infinite collection of trees and regular curves of F;

(3) $u(P)$ is strictly monotone on any cross-section of F;

(4) in every neighborhood of any point P in π there exist points Q for which $u(Q) < u(P)$ and points R for which $u(R) > u(P)$.

It is not hard to show as a consequence of Theorem 1.2-4 that property (3) implies property (4). The function u is pseudo-harmonic as defined in Morse [10, p. 5], and is contoured directly by F.

If u is an allowable function contoured by F, v is an allowable function contoured by G and (F, G) is a conjugate net, then we say the composite function $u + iv$ is an allowable function contoured by (F, G).

Suppose $f(z)$ is defined in a domain D of π and $g(z)$ is defined in a domain E of π. We call $f(z)$ and $g(z)$ *topologically equivalent* [14] if there exists a homeomorphism h of D onto E such that $g(h(z)) = f(z)$ for all z in D.

THEOREM 1.2-9. *Suppose $f(z)$ is an allowable function contoured by a conjugate net (F, G). Then f is topologically equivalent to a function analytic either in π or C_0. Conversely, if g is an analytic function contoured by a conjugate net (F, G) under a homeomorphism h, then g is topologically equivalent under h to an allowable function contoured by (F, G). See* [3, pp. 526-527].

Remark. As a consequence of Theorem 1.2-9 we have the following results. Suppose F contours the real part of an analytic function $f = u + iv$ under a homeomorphism h. Then the level curve family of v is homeomorphic under h^{-1} to a family G complementary to F, the conjugate net (F, G) contours f and f is topologically equivalent under h^{-1} to an allowable function contoured by (F, G).

Suppose $f(z)$ is an allowable function contoured by a conjugate net (F, G). If f is topologically equivalent under a homeomorphism h to an analytic function g, then a path c is a finite asymptotic path of f if and only if $h(c)$ is a finite asymptotic path of g. Thus, for a given family F, in order to show that any analytic function with real part contoured by F must have a finite asymptotic value, it suffices to show that for any family G complementary to F any corresponding allowable function $f = u + iv$ contoured by (F, G) has a finite asymptotic value. For by the remark above, if F contours the real part of an analytic function g, then g is topologically equivalent to an allowable function contoured by a conjugate net (F, G). Conversely, in order to show that there is an analytic function f the real part of which is contoured by F which has no finite asymptotic value, it is sufficient to construct a family G complementary to F and show the existence of an allowable function contoured by (F, G) which has no finite asymptotic value.

2. Finite asymptotic paths and the *b*-curves of *f*.

2.1. *Finite asymptotic paths and b-curves.*

LEMMA 2.1-1. *Let b and c be related b-curves of a branched regular curve family F, and suppose x is a cross-section of F starting at a regular point of c and terminating at a point of a regular curve contained in $\theta(b,c)$. Then b lies entirely on the boundary of $S=F(x)$.*

Proof. By Theorem 1.2-5, the boundary of S consists of regular curves, maximal chains and mixed maximal chains. Since b is a b-curve related to c, either b is a regular curve or a maximal chain, and in the latter case no branched curve of the tree containing b can lie in $\theta(b,c)$. Hence b either lies entirely on the boundary of S or does not intersect the boundary at all. In the latter case, let b' be the boundary of the component of the complement of S which contains b. Since S is contained in $c \cup \theta(b,c)$, $b' \neq c$. But then $b/b'/c$, contrary to the assumption b and c are related b-curves, hence separated by no regular curve or infinite chain of F.

LEMMA 2.1-2. *If (F,G) is a conjugate net and a curve of G tends improperly to infinity relative to F in a given direction, then that curve contains a finite asymptotic path p in the given direction for any allowable function $f=u+iv$ contoured by (F,G).*

Proof. The function u is monotone on p, hence either tends to $\pm\infty$ or is bounded and tends to a finite limit along p. But p tends improperly to infinity, hence the values of u are bounded on p and tend to a finite limit u_0. If $v=v_0$ on p, then $f(z)$ tends to the finite asymptotic value u_0+iv_0 along p.

THEOREM 2.1-1. *Let b and c be related b-curves of a branched regular curve family F. Then any allowable function $f=u+iv$ with real part contoured by F has a finite asymptotic path.*

Proof. Let G be the family complementary to F which contours v, and let P be a regular point of c and $x=G(P)$, the curve of G which contains P. Let x intersect a regular curve s' of $\theta(b,c)$, and let x'' represent the open arc of x between P and $x \cap s'$. Let $S'=F(x'') \cup c$. By Lemma 2.1-1, b lies on the boundary of S'. Let Q be a regular point of b and $y=G(Q)$, and let y intersect a regular curve s of S' in a point A. Since S' is a parallel family, $c/s/s'$. Let y' be the arc of y between Q and A, a be the arc of s between A and $B=x \cap s$, x' be the arc of x'' between P and B and $S=F(x')$. Then $x' \cup a \cup y'$ is a cut of $\theta(b,c)$, separating $\theta(b,c)$ into two disjoint regions C and D. Let D be the region which contains the unbounded portions of s. Since no curve or tree of F can intersect the cross-sections x' and y' more than once and each curve of S intersects x' and y', it follows that all regular curves of S tend to infinity in both directions in D and no branched curve of S can have an endpoint in C. Let b' be the arc of b which lies on the boundary of C. See Figure 2.1.

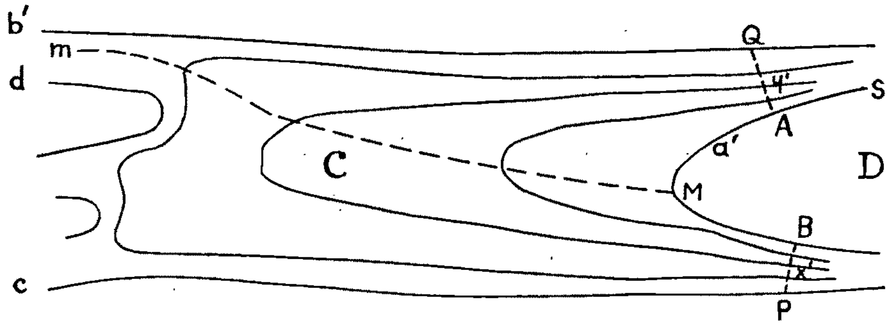

FIGURE 2.1.

Suppose $g \in G$ enters C from a point of b'. The cross-section g of F will intersect a regular curve s'' of S, since the regular curves are dense in π. Since g cannot intersect s'' again, g either must terminate at a branch point in the subregion R of C bounded by x', y', s and s'', intersect x' or y' or intersect s. Note that R is bounded since both s and s' tend to ∞ in both directions in D, hence intersect both x' and y'. But there are no branch points in C and x' and y' are arcs of regular curves of G, hence g does not intersect them. Therefore g intersects s between A and B, i. e. g intersects a. Hence $a \cap G(b')$ is a subarc a' of a with one endpoint A. Let M be the other endpoint of a' and let $m = G(M)$.

Now we show that m tends to infinity in $S \cap C$. Suppose m intersects a boundary element d of $S \cap C$ at a point M'. Then the arc m' of m between a and d is a cut of $S \cap C$: say $S \cap C = C' \cup m' \cup C''$ and suppose b' lies on the boundary of C'. The point M' separates d into two parts, one of which, say d', lies on the boundary of C'. Let N be a point of d' such that $n = G(N)$ is a regular curve. Then by the same argument as above, n intersects a'. But then n separates the connected set $G(b')$, which is impossible. Therefore m intersects no boundary element of C other than s, hence m tends to infinity in $S \cap C$.

Now suppose that $u = u_0$ on c, $v = v_0$ on m and let m be the curve $z = h(t)$, $0 < t < 1$, where we may suppose z tends to infinity along m in $S \cap C$ as t approaches 0. If m intersects curves c', c'' of S at points $h(t')$,

$h(t'')$ respectively, then by Theorem 1.2-7, $c/c'/c''$ if and only if $0 < t' < t''$. Hence as t tends to 0, the corresponding curves of S, $F(h(t))$, intersect the cross-section x' between P and B defined above in points which approach P. Therefore $f(z)$ approaches the finite asymptotic value $u_0 + iv_0$ along $z = h(t)$ as t tends to 0, hence m contains a finite asymptotic path of $f(z)$.

Note. By considering the values of $f(z)$ along b' and a', a similar argument shows that b' is a finite asymptotic path of $f(z)$ and $f(z)$ approaches the value $u_0 + iv_0$ along b'. Also from the above proof we note that b' tends to infinity improperly relative to any complementary family G and the curve m of G tends improperly to infinity relative to F. By Lemma 2.1-2, this implies immediately the existence of a finite asymptotic value of F.

Since by definition b-curves always occur in related pairs, as a consequence of the Remark following Theorem 1.2-9 we have the following corollary.

COROLLARY 2.1-1-1. *If a branched regular curve family F contains a b-curve, then any analytic function with real part contoured by F has a finite asymptotic value.*

COROLLARY 2.1-1-2. *The level curve family of a harmonic polynomial contains no b-curve.*

Proof. Suppose $u(x,y)$ is a harmonic polynomial. Let $v(x,y)$ be a harmonic conjugate. Then $f(z) = u + iv$ is a polynomial, and since the values of a polynomial tend to infinity along any path which tends to infinity, Corollary 2.1-1-2 follows at once from Corollary 2.1-1-1.

THEOREM 2.1-2. *Let c be a regular curve or maximal chain of a branched regular curve family F, and let $D(c)$ be chosen. Suppose $b_1, b_2, b_3, \cdots$ form a (possibly infinite) collection of curves each member of which is contained in $D(c)$ and is a b-curve related to c. Then for any allowable function $f(z) = u + iv$ with real part contoured by F and each b_i, $i = 1, 2, 3, \cdots$, there corresponds a finite asymptotic path g_i such that if $i \neq j$, $i, j = 1, 2, 3, \cdots$, then the corresponding finite asymptotic values of $f(z)$ approached along g_i and g_j respectively are distinct.*

Proof. Let V designate the collection of curves $c, b_1, b_2, b_3, \cdots$. Since b_i is contained in $D(c)$ and is related to c for each $i = 1, 2, 3, \cdots$, V is a cyclic set. Let the curve family G and the point P of c be given as in the proof of Theorem 2.1-1, and let $x = G(P)$ intersect a regular curve s of $\theta(V)$. Let x' be the arc of x between P and s, and let $S = F(x')$. Let K_n be a disk with center P and radius n, $n = 1, 2, 3, \cdots$. It follows from Theorem 1.2-3 by an argument similar to that in [8, p. 175], that at most a finite number of elements of V intersect K_n for each n. Suppose n is fixed. Then as in the proof of Theorem 2.1-1, for each element b_i of V which intersects K_n a regular curve s_i of S can be found which is accessible by an arc of a regular curve of G from a point of b_i. By an extension of Theorem 1.2-7, one of these curves s_i, call it $s(n)$, will separate the other s_i from c and will be accessible by elements of G from each of the curves b_i which intersect K_n. Note that in general $s(n)$ may not be accessible from elements of V which do not intersect K_n; in particular the existence of a branched curve of S may interfere.

Now consider the elements of V which intersect K_{n+1}. There exists a regular curve $s(n+1)$ of S such that $c/s(n+1)/s(n)$ and $s(n+1)$ is accessible from each element which intersects K_{n+1}. Note that a cross-section from b_i to $s(n)$ will also contain a cross-section from b_i to $s(n+1)$.

As in the proof of Theorem 2.1-1, corresponding to each element b_i of V which intersects K_n there will be an arc $a_i(n)$ of $s(n)$ consisting of the intersections with $s(n)$ of the curves of G which intersect a half-infinite subarc of b_i, $a_i(n)$ having endpoints $A_i(n)$ and $M_i(n)$, such that $G(M_i(n))$ contains a finite asymptotic path of any allowable function $f(z) = u + iv$ contoured by (F, G). Since v is monotone on $s(n)$ for given n, the finite asymptotic values approached along the curves $G(M_i(n))$ corresponding to the elements of V which intersect K_n are all distinct complex numbers. If $G(M_i(n))$ has been chosen for $s(n)$ accessible from b_i, then $G(M_i(n))$ also intersects $s(n+r)$, $r = 1, 2, \cdots$, in the points $M_i(n+r)$. Hence $G(M_i(n+r))$ is independent of r. Let $g_i = G(M_i(n))$. Since any element of V will intersect K_n for n sufficiently large, a finite asymptotic path g_i is associated with each element b_i of V. Also, if b_i and b_j are any two elements of V, then there will be an integer n such that both b_i and b_j intersect K_n, and as noted above the finite asymptotic values approached along the corresponding finite asymptotic paths g_i and g_j are distinct.

Note. As in Theorem 2.1-1 it can be shown that each of the b_i contains a finite asymptotic path along which $f(z)$ approaches the same value as along the corresponding path g_i.

2.2. *Cyclic sets and b-curves.* Since c and the related b-curves b_i in Theorem 2.1-2 form a cyclic set, this would indicate that there might be some relation between the existence of a cyclic set and the existence of finite asymptotic paths. That a cyclic set does not necessarily imply the existence of a finite asymptotic path is shown by the example of the level curve family of $u = Re(z^2)$, where for fixed $c_1 > 0$ and $c_2 < 0$ the curves $u = c_1$ and $u = c_2$ form a cyclic set V while $f(z) = z^2$ has no finite asymptotic value. In this example $\theta(V)$ contains the branch point $z = 0$. It turns out this is crucial; in fact we have the following theorem.

THEOREM 2.2-1. *If F is a branched regular curve family and V is a cyclic set consisting of three or more elements of F such that $\theta(V)$ contains no branched curve of F, then $\theta(V) \cup V$ contains at least two related b-curves.*

Proof. Let V contain the elements b, c and d, P be a regular point of b and x be a cross-section of F which tends to infinity in the side $D(b)$ which contains c and d. Theorem 1.2-4 guarantees the existence of such a cross-section. Let $x' = F(x) \cap \theta(V)$, $S = F(x')$ and let c' (d') be the boundary of the component of the complement of S which contains c (d). Suppose $x' \cup P$ is represented by $z = h(t)$, $0 \leqq t < t_1 \leqq 1$, where $P = h(0)$. Suppose x intersects d, so $h(t_1)$ is a point of d. We distinguish four possible cases.

Case 1. For all s in S, $b/s/c'$. Then $|c', s, d|$ for all s in S, and c' and d are related b-curves since $d = d'$.

Case 2. There exists t_0, $0 < t_0 < t_1$, such that if $0 < t \leqq t_0$, then $b/F(h(t))/c$ while if $t_0 < t < t_1$, then $|b, F(h(t)), c|$. Then $F(h(t_0))$ and c' are related b-curves.

Case 3. There exists t_0, $0 < t_0 < t_1$, such that $b/F(h(t))/c'$ whenever $0 < t < t_0$, while $|b, F(h(t)), c'|$ if $t_0 \leqq t < t_1$. Then $F(h(t_0))$ is again a b-curve related to c'.

Case 4. For all s in S, $|b, s, c'|$. Then b and c' are related b-curves.

Suppose $0 < t < t' < t_1$. Then it is impossible for both $|b, F(h(t)), c'|$ and $b/F(h(t'))/c'$ to hold. Hence the above cases are the only ones possible when x intersects d, and c' is always a b-curve related to b, d or an element of S.

The proof is similar when x intersects another boundary element of V or tends to infinity in $\theta(V)$.

A somewhat complicated induction proof [13] shows that whenever V contains $n \geqq 3$ elements and $\theta(V)$ contains no branched curve, then there exists a set of at least $n - 1$ related b-curves in $\theta(V) \cup V$. However, some or all of the corresponding finite asymptotic paths may be equivalent, as is shown by the example of the set of level curves $Re(e^z) = c$ for fixed $c \neq 0$

which is an infinite cyclic set of a regular curve family with no branched curves contouring the real part of an analytic function with only one finite asymptotic value.

Theorem 2.2-1 has no exact converse, as the example illustrated in Figure 2.2 shows. In Figure 2.2 the branch points recede to infinity to the right. The family is a parallel family in each component of the complement of $\theta(b,c)$. The curves b and c are related b-curves. Note that the trees are not isolated. However we do have the following result.

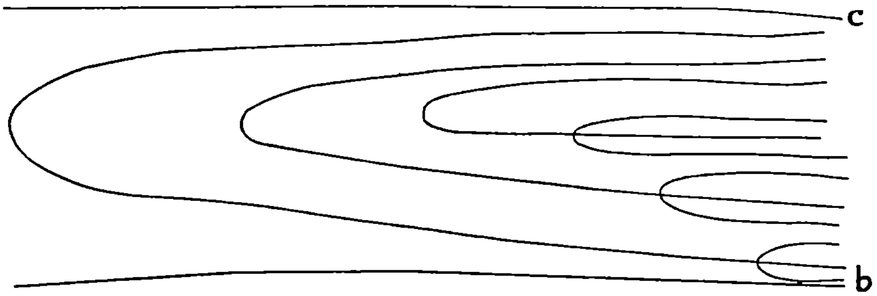

FIGURE 2.2.

THEOREM 2.2-2. *If F is a branched regular curve family which contains a b-curve and the trees of F are isolated, then there exists a cyclic set V containing three or more elements such that $\theta(V)$ contains no branched curve.*

Proof. Let b be a b-curve of F and suppose c is a related b-curve. Let x be a cross-section of F starting at a regular point P of b and terminating at a point Q of a regular curve r of $\theta(b,c)$. Then by Lemma 2.1-1, b and c lie entirely on the boundary of $S = F(x)$. Let x be the curve $z = h(t)$, $0 \leqq t \leqq 1$, where $h(0) = P$ and $h(1) = Q$. Since the trees of F are isolated, at most a finite number of branched curves can intersect x. Hence there exists a t', $0 < t' \leqq 1$, such that for $0 < t \leqq t'$, $F(h(t))$ is a regular curve. Let x' be the subarc $z = h(t)$, $0 \leqq t \leqq t'$, and $S' = F(x')$. Then $b, c, F(h(t'))$ and the other elements on the boundary of S' form the desired cyclic set by Theorem 1.2-6.

COROLLARY 2.2-2-1. *If the trees of F are isolated, then F contains a b-curve if and only if F contains a cyclic set V containing three or more elements such that $\theta(V)$ contains no branched curve.*

We conclude this section by observing that as a consequence of Lemma 2.1-2 we can formulate another sufficient condition that F contain a finite asymptotic path for any allowable function with real part contoured by F, namely: Each family G complementary to F contains a curve which tends improperly to infinity relative to F. The existence of such a curve is implied by the existence of a b-curve in F.

3. Sufficient condition that F contour an allowable function with no finite asymptotic value. In this section we prove the principal sufficiency theorem: If F contains no b-curves and the trees of F are isolated, then F contours the real part of an allowable function with no finite asymptotic value. An interesting consequence of this result is that the level curve families of harmonic polynomials are characterized by the properties that each family contains only a finite number of branch points and no b-curves.

Throughout this section we assume that F is a branched regular curve family filling the plane π, F contains no b-curve and the trees of F are isolated, i. e. at most a finite number of trees intersect any given bounded subset of π. We first partially order the trees of F by separation and orient the regions determined by the trees. Then we define an allowable function u contoured by F and show the existence of a family G complementary to F and an allowable function v contoured by G such that the allowable function $u+iv$ contoured by (F, G) has no finite asymptotic value.

3.1. *The regions of* $\pi - \Sigma$. Let Σ denote the set of trees of F and let $H = \pi - \Sigma$.

THEOREM 3.1-1. *The set H is the union of open disjoint regions each of which has at most two elements of F on its boundary.*

Proof. Let P be an element of H and N a bounded neighborhood of P. The point P, being a regular point, is not a limit point of any branched curve, and since the trees of F are isolated, Theorem 1.2-3 implies that at most a finite number of branched curves can intersect N. Hence there is a neighborhood of P contained in N which lies entirely in H and so H is an open set. Let Q be a component of H. Since each boundary element of Q separates π, the boundary of Q forms a cyclic set V. Since Q contains no branched curves, if V contains three or more elements, then by Theorem 2.2-1 F contains a b-curve. Hence V contains at most two elements.

THEOREM 3.1-2. *Let Q be a component of H. Then $F[Q]$ is a parallel family, and if Q has two boundary elements, then each curve of $F[Q]$ separates them. Also, if c is a boundary element of Q, $D(c)$ contains Q and x is a cross-section of F which tends properly to infinity in $D(c)$, then* $F[Q] = F(x \cap Q)$.

Proof. Let c be a boundary element of Q and choose $D(c)$ to contain Q. Let x be a cross-section of F starting at a regular point P of c and tending properly to infinity in $D(c)$, and let $S = F(x \cap Q)$. Suppose Q has only the one boundary element c. Then $Q = D(c)$ and contains $F(x)$. Suppose q is a curve of $F[Q]$ not in $F(x)$. If all curves of S separate q and c, then x does not tend properly to infinity. If some curve s of S does not separate q and c, then q, s and c form a cyclic set V such that $\theta(V)$ contains no branched curve. But then F contains a b-curve. In either case a contradiction is reached, so $F(x) = F[Q]$.

Suppose Q has a second boundary element d. If some curve of $F(x \cap Q)$ does not separate c and d, then as above F contains a b-curve. Hence x intersects d and again either F contains a b-curve or $F(x \cap Q) = F[Q]$.

3.2. *Adjacent sequences of trees and regions.* Throughout the remainder of Section 3 a *region* will mean a region of H unless otherwise specified. Two regions will be said to be *adjacent regions* if they have a common branched curve on their boundaries. An *adjacent sequence of regions* is a sequence of regions $R_1, R_2, \cdots, R_n$ such that R_i and R_{i+1} are adjacent, $i = 1, 2, \cdots, n-1$. If an adjacent sequence of regions contains the regions R and S, we say R is *connected* to S by the sequence. An adjacent sequence of regions connecting two regions R and S will be said to be a *minimal sequence* of n regions if it contains exactly n regions and no adjacent sequence of fewer than n regions connects R and S.

A region will be said to be *adjacent to a tree* T if there is a maximal chain of T belonging to its boundary; the tree will also be said to be adjacent to the region. Two trees will be said to be *adjacent trees* if they are adjacent to a common region. Let T_0 be a fixed tree of F. If T and T' are trees and T separates T_0 and T', then we write $T < T'$. We define $T_0 < T$ for all $T \neq T_0$. The relation $<$ gives a partial ordering of the trees of F relative to T_0. An *adjacent sequence of trees* connecting T_0 and T is a set of trees $T_1, T_2, \cdots, T_n$ such that $T_1 = T_0$, $T_n = T$, T_i and T_{i+1} are adjacent trees and $T_i < T_{i+1}$, $i = 1, 2, \cdots, n-1$.

THEOREM 3.2-1. *Let T_0 be a fixed tree of F. If T is any tree of F, then there is a unique adjacent sequence of trees connecting T_0 and T.*

Proof. Suppose $T' < T$ and $T'' < T$, $T' \neq T_0 \neq T''$. The complement of T' is the union of disjoint regions of π. Let U be the region containing T_0 and V the region containing T. Since $T' < T$, U and V are distinct. If T'' lies in U, then, since $T'' < T$; $T'' < T'$. If T'' lies in V, then $T' < T''$.

If T'' lies neither in U nor V, then there is a path in U from T_0 to a point P of T', there is a path in V from T to a point Q of T' and from the definition of a tree there is a path along a finite chain of T' from P to Q. Hence $T'' < T$ is false, contrary to hypothesis. Therefore the trees separating T_0 and T can be simply ordered by $<$. Since the trees of F are isolated, there can be but a finite number of them separating T_0 and T and simply ordering them by $<$, we get a unique adjacent sequence connecting T_0 and T.

Definition 3.2-1. Let $n(T)$ represent the number of trees in the adjacent sequence of trees connecting T_0 and T. I.e. $n(T) - 2$ represents the number of trees separating T_0 and T, $n \geqq 2$.

THEOREM 3.2-2. *If R and S are any two regions of H, then there is a minimal sequence of regions of H connecting R and S.*

Proof. Let A be the union of all the regions of H which can be connected to a fixed region R by an adjacent sequence. Let A' be the interior of $\bar{A}$, where $\bar{A}$ is the closure of A, and $B = \pi - A'$. Let $P \in B$. If P is not on a tree, then P is contained in a region of H contained in B. If P is a regular point of a tree, then by Theorem 1.2-1, P is on the boundary of exactly two regions, V and W, of H. If V is in A, then V is connected to R by an adjacent sequence ρ, hence W is connected to R by the sequence ρ, V, so W is also in A. But then $P \in A'$. Similarly if W is contained in A, $P \in A'$. Therefore $P \in \operatorname{Int}(\bar{V} \cup \bar{W}) \subset B$. If P is a branch point, then by Theorem 1.2-1 there are exactly $2n$ regions determined by the $2n$ maximal chains which contain P. If any one of these regions is contained in A, then all of them are contained in A, since in going around P in a clockwise direction any two of these regions are clearly connected by an adjacent sequence of $2n$ or fewer of these regions. Therefore P again is contained in an open set contained in B. Hence B is an open set. But then π is the union of two disjoint open sets, so, since A contains R, hence is nonempty, B is empty. Therefore each region of H can be connected to a fixed region of H by an adjacent sequence. Among the adjacent sequences connecting two regions there will be one or more minimal sequences.

THEOREM 3.2-3. *If R and S are distinct regions of H each connected to a given region R_0 by a minimal sequence of n regions, $n \geqq 2$, then R and S are not adjacent.*

Proof. Since each component of the boundary of a region of H is a maximal chain and each branch point lies on the boundary of at least four regions of H, no three regions can be adjacent to each other. Therefore if $n=2$, then R and S are adjacent to R_0, hence not adjacent to each other.

Suppose the theorem is true for sequences containing $n-1$ regions or less and suppose R and S are regions of H connected to R_0 by respective minimal sequences $\rho\colon R_0=R_1, R_2, \cdots, R_n=R$ and $\sigma\colon R_0=S_1, S_2, \cdots, S_n=S$, $n>2$. If for some $i\geqq 2$, R_i and S_i coincide, then by the induction hypothesis applied to the sequences $R_i, R_{i+1}, \cdots, R$ and $R_i=S_i, S_{i+1}, \cdots, S$, R and S are not adjacent.

Suppose R_i and S_i are distinct for all $i\geqq 2$. Let R_2 be adjacent to R_0 along a curve b and S_2 be adjacent to R_0 along a curve c. If b and c belong to distinct boundary chains of R_0, then R_0 separates R and S. Suppose b and c belong to a boundary chain k of R_0 and are separated on k by a curve d. Then the region W adjacent to R_0 along d belongs neither to ρ nor σ, since ρ and σ are minimal sequences. Therefore $\overline{R_0\cup W}$ separates R and S. Suppose b and c have a common endpoint B. If W is one of the regions of H with B on its boundary and W is not an element of ρ or σ, then $\overline{R_0\cup W}$ separates R and S. Suppose all regions of H with B on their boundary are in ρ or σ. Then there will be a branched curve of $St(B)$ on the boundary of both a region of ρ and a region of σ. Since there are an even number of regions with B on their boundary of which R_0 is one, R_i and S_i are distinct for $i>2$ and ρ and σ are minimal sequences, the subscripts of these adjacent regions will differ by exactly one, say the adjacent regions are R_i and S_{i+1}, $i\geqq 2$. Then the sequences $R_i, R_{i+1}, \cdots, R$ and $R_i, S_{i+1}, \cdots, S$ satisfy the induction hypotheses, so R and S are not adjacent in this case either.

In order to define the functions u and v and the complementary family G it will be convenient to define a positive direction on each chain of F and an orientation for each region. Throughout the remainder of this section R_0 will denote a fixed region adjacent to the fixed tree T_0.

We now associate with each region a sequence of numbers.

Definition 3.2-2. Let R be a region. Then $m(R)$ will denote the minimum of the numbers $n(T)$ for all trees T adjacent to R.

Let R_0 have the sequence (1). Suppose S is a region adjacent to R_0. Then S will have a sequence $(1,\epsilon)$, where $\epsilon=1$ if $m(S)>1$, $\epsilon=-1$ if $m(S)=1$. Suppose sequences have been chosen for all regions connected to R_0 by minimal sequences of $n-1$ or fewer regions, one sequence of numbers for each minimal sequence of regions. Let S be a region connected to R_0 by a minimal sequence of regions σ with n terms, and suppose R is the region of σ adjacent to S. Let ρ be the sequence of numbers corresponding to the subsequence of $n-1$ regions of σ connecting R_0 and R. If the last term in ρ is 1, let the sequence corresponding to S be $(\rho, 1)$ if $m(R) < m(S)$ and $(\rho, -1)$ if $m(R) = m(S)$. If the last term of ρ is -1, let the sequence corresponding to S be $(\rho, -1)$ if $m(R) < m(S)$ and $(\rho, 1)$ if $m(R) = m(S)$. See Figure 3.1.

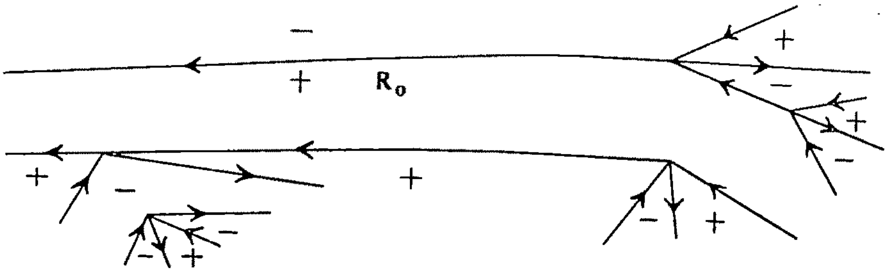

FIGURE 3.1.

Since by Theorem 3.1-1 no region can be adjacent to three trees, it follows that if R is a region adjacent to two trees T and T', then either $T < T'$ or $T' < T$, and if $n(T) < n(T')$, then $T < T'$. As a consequence, if R and S are adjacent regions and $m(R) < m(S)$, then either $R = R_0$ or R separates R_0 and S. Conversely, it is clear that if R separates R_0 and S, then, since R_0 is adjacent to T_0, $m(R) < m(S)$. Hence if S_i and S_{i+1} are two consecutive regions in a minimal sequence σ connecting R_0 and S, $i > 1$, and ρ is the corresponding sequence of numbers, then the terms in ρ corresponding to S_i and S_{i+1} will have like signs if and only if S_i separates R_0 and S_{i+1}. If S_i separates R_0 and S_{i+1}, then S_i separates R_0 and S. Hence S_i must occur in the same place in any other minimal sequence of regions σ' connecting R_0 and S. Furthermore if S'_{i+1} is the region of σ' following S_i, then $m(S_i) < m(S'_{i+1})$, so that in the sequence of numbers ρ' corresponding to σ', the terms corresponding to S_i and S'_{i+1} also have like signs. Since all sequences start with 1, it follows that ρ and ρ' coincide and there is a unique sequence associated with each region.

Definition 3.2-3. Let $\gamma(R)$ be the sign of the last term in the sequence for R.

Definition 3.2-4. Let $n(R)$ denote the number of terms in a minimal sequence connecting R_0 and R.

3.3. *Orientation of the maximal chains of F.* Let R be a region. If $\gamma(R)$ is $+$, give R a positive (counterclockwise) orientation, while if $\gamma(R)$ is $-$, give R a negative (clockwise) orientation. Suppose k is a maximal chain of a tree T. It will lie entirely on the boundary of one and only one region R. If $m(R)=n(T)$, then define the positive direction on k so it coincides with the orientation of R. In more detail, let B be a branch point of k and b and c be the elements of $St(B)$ which lie on k. Suppose R is positively oriented. Then choose the positive direction on k so that going along k in the positive direction, if b is reached first before c, then c, b are counterclockwise adjacent, while if c is reached first, then b, c are counterclockwise adjacent. If R has a negative orientation, then define the positive direction on k so that going along k in the positive direction, if b is reached first, then c, b are clockwise adjacent, while if c is reached first, then b, c are clockwise adjacent. If $m(R)<n(T)$, then define the positive direction on k to be the opposite to the orientation of R.

THEOREM 3.3-1. *The maximal chains of F are consistently oriented.*

Proof. Let k be a maximal chain and b a curve of k. By Theorem 1.2-1, b is contained in exactly two maximal chains k and k' lying on the boundaries of regions R and S respectively. Since R and S are adjacent, by Theorem 3.2-3, $n(R)$ and $n(S)$ differ by exactly one.. Suppose $n(R)<n(S)$, and suppose R and S have the same orientation. Then $m(R)<m(S)$, so $m(R)<n(T)$, T the tree on which k and k' lie, and k receives a positive direction opposite to the orientation of R while k' receives a positive direction coinciding with the orientation of S. Hence b receives the same positive direction considered as a curve of k or k'. If R and S have opposite orientations, then $m(R)=m(S)$, and $m(R)=n(T)=m(S)$. Therefore k receives a positive direction coinciding with the orientation of R while k' receives a positive direction coinciding with the orientation of S. Since R and S have opposite orientations, b receives the same positive direction considered as a curve of k or k'. Hence the maximal chains of F have been consistently oriented.

3.4. *The definition of $u(P)$.* We now define an allowable function u contoured by F such that $|u|$ tends to ∞ on any cross-section of F which tends properly to infinity.

Let $T\neq T_0$ be a tree of F. Suppose T' is the tree adjacent to T in the adjacent sequence of trees connecting T_0 and T and R is the region adjacent to both T and T'. If $\gamma(R)>0$, let $u=n(T)$ on T, while if $\gamma(R)<0$, let $u=-n(T)$ on T. Let $u=1$ on T_0.

Suppose R is a region with a maximal chain k of a tree T on its boundary. Choose $D(k)$ to contain R and let x: $z=f(t)$, $0\leqq t<1$, be a cross-section tending properly to infinity in $D(k)$ from a point P of k. Suppose $u=u_0$ on k. If k is the only boundary chain of R, then define

$$u=|u_0|(1+\tan(\pi/2)t)\gamma(R)$$

on the curve $F(f(t))$, $0<t<1$. Note that since $n(T)\geqq 1$, $u_0\neq 0$. If R has a second boundary chain, then x intersects it at some point $f(t_0)$, $0<t_0<1$, as shown in the proof of Theorem 3.1-2. Suppose $u=u_1$ at $f(t_0)$. Then we define $u=u_0+(t/t_0)(u_1-u_0)$ on $F(f(t))$, $0<t<t_0$. Defining u in similar fashion on all the regions, Theorem 3.1-2 and the discussion in Section 1.2.7 guarantee that u is defined and continuous on all of F and is contoured by F.

Note 1. If R has only one boundary chain, then u tends to $\gamma(R)\infty$ as t tends to 1.

Note 2. If R has two boundary chains which are subsets of trees T and T' respectively such that $n(T)<n(T')$, then if R has a positive orientation, $u_1=|u_0|+1>u_0$ and u increases on x as t tends to t_0, while if R has a negative orientation, $u_1=-|u_0|-1<u_0$ and u decreases on x as t tends to t_0.

We now verify that u is an allowable function contoured by F. See Section 1.2.8. Property (1) follows from the definition. Since the trees and regions are countably infinite at most and u assumes a given value on at most one curve in each region, Property (2) is also satisfied. Suppose x is a cross-section of F and x is contained in a region R. Since u is continuous on x, x intersects each curve of F at most once and u assumes a given value on at most one curve of $F[R]$, u must be monotone on x. Suppose x intersects two adjacent regions, R and S. Then u is monotone on $x\cap R$ and $x\cap S$. Suppose R and S both have positive orientation and $n(R)<n(S)$. Then $m(R)<m(S)$ and either $R=R_0$ or R separates R_0 and S. In either case there is a tree T adjacent to R and S and a tree T' adjacent to R such that $n(T')<n(T)$, and by note 2 above u increases on x as x approaches T. Since $\gamma(R)>0$, $u=u_0>0$ on T, and since $\gamma(S)>0$, u increases on x as x recedes from T as a consequence of notes 1 and 2 above. Hence u is monotone on $x'=(x\cap\bar{R})\cup(x\cap\bar{S})$. A similar argument shows u is monotone on x' if $\gamma(R)$ and $\gamma(S)$ are both negative.

Suppose R and S have opposite orientations. Then $m(R)=m(S)$, and again it follows from the definition of u and notes 1 and 2 that u is monotone on x'. Thus u is monotone on each section x' of x which intersects two adjacent regions, and therefore u is monotone on all of x.

Suppose x is a cross-section of F which tends properly to infinity. If x intersects an infinite number of trees, then these trees can be simply ordered by $<$, and $|u|$ tends to ∞ on x. If x intersects only a finite number of trees, then x tends to infinity in a region R. It follows from Theorem 3.1-2 that R has but one boundary chain k and, after extending x to k if necessary, $F[R] = F(x \cap R)$. From the definition of u in R, $|u|$ tends to ∞ on $x \cap R$.

3.5. *The definition and properties of G and v.* In this section we prove the following theorem.

THEOREM 3.5-1. *There exists a family G complementary to F such that each regular curve or infinite chain of G tends properly to infinity in each direction relative to F, and there exists an allowable function v contoured by G such that $|v|$ tends to ∞ on each element of F and to no finite limit on any infinite chain of F.*

3.5.1. *The definition of v and G.* First we define v on the trees of F. Let T be a tree of F and B_0 a fixed branch point of T. Let $v = 0$ at B_0. Suppose B is a branch point in $St(B_0) - B_0$. Then the branched curve B_0B belongs to exactly two maximal chains k and k' of T which have been given a consistent orientation in Section 3.3. If the direction from B_0 to B is positive along k, let $v = 1$ at B; if B is in the negative direction along k from B_0, let $v = -1$ at B. In general if v has been assigned the value n at a branch point B of T, then let $v = n + 1$ at a branch point B' of $St(B) - B$ if the direction from B to B' is positive, $v = n - 1$ at B' if the direction from B to B' is negative. Since there is a unique finite chain connecting any two branch points of T, v will be defined and single valued at each branch point of T.

Suppose b is a branched curve $z = f(t)$, $0 < t < 1$, and b has two end points, $B = \lim_{t \to 0} f(t)$ and $B' = \lim_{t \to 1} f(t)$ on T. If $v = n$ at B and $n + 1$ $(n - 1)$ at B', then define $v = n + t$ $(v = n - t)$ at $f(t)$, $0 < t < 1$. If b has only one end point B, where $v = n$, let

$$v = n + \tan(\pi/2)t \quad (v = n - \tan(\pi/2)t)$$

if the direction from B to ∞ along b is positive (negative). Define v in similar fashion on each tree of F. Then v is continuous and monotone on each maximal chain k of F, and $|v|$ tends to ∞ in each direction on k.

Let R be a region and let k be a maximal chain lying on the boundary of R. Choose $D(k)$ to contain R and let x be a cross-section which tends properly to infinity in $D(k)$ from a regular point on k. Then as a consequence of Theorem 3.1-2 and the results in Section 1.2.7, $k \cup R$ can be mapped by a homeomorphism h onto the family of lines $y = c$, $0 \leqq c < 1$, in the xy-plane, k going onto the line $y = 0$. Let $R' = \bar{R}$ minus the set of branch points of F. If R has no other boundary element, then define the curves of G in R' to be the intersection with R' of the inverse images under h of the line segments σ_c: $x = c$, $0 \leqq y < 1$, $-\infty < c < \infty$, and define v on $h^{-1}(\sigma_c)$ to have the value of v at $h^{-1}(c, 0)$ on k. If R is adjacent to another maximal chain m, then x must intersect m, and $k \cup R \cup M$ can be mapped by a homeomorphism h onto the family of lines $y = c$, $0 \leqq c \leqq 1$, k going onto the line $y = 0$, m going onto the line $y = 1$. Connect by straight line segments the images under h of the points P and P', on k and m respectively, where $v(P) = v(P')$. Since by the definition of the orientations of the maximal chains of F given in Section 3.3 both k and m have the same positive direction, these line segments will not intersect each other. Define G in R' to be the inverse images under h of these line segments. Since $|v|$ tends to ∞ in both directions and is monotone on both k and m, all the points of k and m will be connected in a one-to-one fashion by curves of $G[R']$. Define v on $G[R']$ as above. Let two arcs or half-open curves of G be part of the same curve of G if they have a common endpoint which is a regular point of F.

3.5.2. *Verification that G is complementary to F.* From the definition, each curve of G is a cross-section of F and G is regular at each point of H. Let P be a point of a tree T. By Theorem 1.2-1, P belongs to exactly $2n$ maximal chains, $k_1, k_2, \cdots, k_{2n}$, where $n = 1$ if P is a regular point. Suppose the maximal chains have been labeled so that k_i and k_{i+1} have the branched curve b_i in common and b_i, b_{i+1} are clockwise adjacent, $i \equiv 1, 2, \cdots, 2n$ (mod $2n$).

We now define a mapping of a deleted neighborhood of P onto an n-fold covering of $0 < |z| \leqq 1$. Each chain k_i lies on the boundary of exactly one region R_i, and there is an associated homeomorphism h_i defined in Section 3.5.1 of $k_i \cup R_i$ onto the strip $0 \leqq y < 1$, $-\infty < x < \infty$, with curves of $G[k_i \cup R_i]$ going onto line segments connecting points on $y = 0$ and $y = 1$. Let P_i be a point of b_i and $s_i = h_i^{-1}$ $(y = c)$, $0 < c < 1$, be a curve of $F[R_i]$ for each i. Let g_i be the arc of $G(P_i)$ between P_i and s_i, g_i' be the arc of $G(P_i)$ between P_i and s_{i+1}, $i \equiv 1, 2, \cdots, 2n$ (mod $2n$), and let R_i' be the simply connected region bounded by k_i, g_i, g'_{i-1} and s_i. Then h_i maps R_i' onto a trapezoidal region bounded by $y = 0$, $y = c$, $h(g_i)$ and $h(g'_{i-1})$. If necessary modify h_i by a reflection in the y-axis so that the positive direction on k_i corresponds to the positive direction on the x-axis under h_i. Now we map $h_i(\bar{R}_i')$ onto a rectangular region by a homeomorphism h_i'. Let $h_i(P) = (x_0, 0)$, $h_i(P_{i-1}) = (x_1, 0)$, $h_i(P_i) = (x_2, 0)$ and for each point (x, y) of $h_i(\bar{R}_i')$ let $(x'', 0)$ be the endpoint on the x-axis of the segment through (x, y) which corresponds under h_i^{-1} to a curve of $G[R_i]$. Define $h_i'(x, y)$ to be the point (x', y'), where

$$x' = (x_0 - x'')/(x_1 - x_0),\ x'' \text{ between } x_0 \text{ and } x_1,$$
$$x' = (x'' - x_0)/(x_2 - x_0),\ x'' \text{ between } x_0 \text{ and } x_2,$$

if the direction from P_{i-1} to P_i on k_i is positive; otherwise

$$x' = (x_0 - x'')/(x_0 - x_1),\ x'' \text{ between } x_0 \text{ and } x_1,$$
$$x' = (x'' - x_0)/(x_0 - x_2),\ x'' \text{ between } x_0 \text{ and } x_2.$$

Let $y' = y/c$ if i is odd, $y' = -y/c$ if i is even. Now follow h_i' by a mapping h_i'' such that $h_i''(x, y) = (x', y')$, where $y' = y$ and x' is the x-coordinate of $h'_{i-1}h_{i-1}h_i^{-1}h'^{-1}_i(x, 0)$ for x between 0 and $h_i'h_i(P_{i-1})$, $x' = x$ for x between 0 and $h_i'h_i(P_i)$, $i \equiv 1, 2, \cdots, 2n \pmod{2n}$. Let $f_i = h_i''h_i'h_i$. Since the maximal chains of F have been consistently oriented, f_i is a homeomorphism which coincides with f_{i+1} along the arc of b_i from P to P_i.

Let V_i be the image under f_i^{-1} of the upper half of $|z| \leqq 1$ if i is odd, the lower half of $|z| \leqq 1$ if i is even. Then the homeomorphisms $f_1, f_2, \cdots, f_{2n}$ together with the regions $V_1, V_2, \cdots, V_{2n}$ define an n-fold covering of $0 < |z| \leqq 1$ by a deleted neighborhood of P. The subsequent map $z^{1/n}$ will then complete the mapping of $V = V_1 \cup V_2 \cup \cdots \cup V_{2n}$ onto $0 \leqq |z| \leqq 1$, the curves of $G[V]$ going into arcs of $Re(z^n) = \text{constant}$, P going into 0. Hence P is a branch point of G of order n if $n > 1$, a regular point of G if $n = 1$. Therefore G is a branched regular curve family complementary to F.

3.5.3. *Some properties of G and v.* First we verify that each regular curve or infinite chain c of G tends properly to infinity relative to F. Since the trees of F are isolated, if c intersects an infinite number of trees of F in a given direction, then c tends properly to infinity in that direction. If c intersects but a finite number of trees of F in a given direction, then, since it can intersect a given tree at most once, it tends to infinity in a region R. By the definition of G, R has but one boundary chain, and c tends properly to infinity in R.

Next we verify that $|v|$ tends to ∞ in each direction on each element of F and tends to no finite limit on any infinite chain of F. By definition, $|v|$ tends to ∞ in each direction on a maximal chain of F. If c is a regular curve of F, then c lies in a region R and every curve of G which intersects a given boundary chain k of R also intersects c. Hence $|v|$ tends to ∞ on c in each direction. Suppose c is an infinite chain of F. Since v differs by ± 1 at the endpoints of a branched curve, v tends to no finite limit on a chain on which the branch points tend to ∞ in both directions. If the branch points on c do not tend to infinity in a certain direction, then beyond a certain point c consists of a single branched curve b and $|v|$ tends to ∞ on b, hence on c.

Finally we show v is an allowable function contoured by G. Properties (1) and (2) follow from the definition of v. Suppose x is a cross-section of G contained in a region R. Each curve of G which intersects x also intersects a boundary chain k of R. Since v is monotone on k, it is also monotone on x. Suppose x is a cross-section of G which intersects two adjacent regions R and S. Let R and S have the common boundary curve b. Since x is a cross-section of G and G is complementary to F, $x \cap b$ is a regular point P of F. Let V be the neighborhood of P described in Section 3.5.2 $(n=1)$. Then each curve of G which intersects $x \cap V$ also intersects b and since v is monotone on b, v is monotone on $x \cap V$. Hence v is monotone on x.

3.6. *The main sufficiency theorem.*

THEOREM 3.6-1. *Suppose F is a branched regular curve family, the trees of F are isolated and F contains no b-curve. Then there exists an analytic function the real part of which is contoured by F which has no finite asymptotic value.*

Proof. Let u, v and G be defined as in Sections 3.4 and 3.5. Suppose c is a path which tends to infinity from a fixed point P. If c intersects an infinite number of trees of F, then $\limsup |u| = \infty$ on c.

Suppose c intersects one tree T of F in an unbounded set of points. Let M and N be two points of intersection of c and T. There is a unique chain k of T connecting M and N. If M and N are both points or endpoints of a single branched curve b, then, since v is monotone on b and continuous on c, v assumes every value on c it does on b between M and N. Suppose k contains branch points between M and N. In going from M to N on k, suppose B_1 is the first branch point encountered beyond M and let M and B_1 lie on the closure of the branched curve b_1. Let m_1 be the mixed maximal chain through B_1 determined by b_1. The chain m_1 separates M and N, so c intersects m_1 in a point P_1. The chain connecting M and P_1 is part of a maximal chain k_1. Since v is monotone on k_1 and continuous on c,

v assumes every value on c between M and P_1 that it assumes on k between M and B_1. Let B_2 be the next branch point beyond B_1 on k and P_2 a point in which c intersects the mixed maximal chain through B_2 determined by the branched curve B_1B_2. Again it follows that v assumes every value on c between P_1 and P_2 that it does on k between B_1 and B_2. Continuing in this manner, we conclude that v assumes every value on c between M and N that it does on the chain k between M and N. In particular if there are at least two branch points on k between M and N, then v assumes values on c between M and N which differ by an integer. Hence if the set S of branch points on chains connecting the intersections of c and T is unbounded, v tends to no finite limit on c. If S is bounded, then c will intersect one branched curve b in an unbounded set of points, and in this case $\limsup |v|$ will tend to ∞ on c.

If c intersects no tree of F in an unbounded set of points and c intersects at most a finite number of trees of F, then c tends to infinity in a region R. If R has only one boundary chain and c intersects each curve of $F[R]$, then $\limsup |u| = \infty$ on c. Suppose c does not intersect a curve s of $F[R]$, where R may have two boundary chains. Then c tends to infinity in the region bounded by s and a boundary chain k of R. Let c intersect a regular curve g of $G[R]$ and let $Q = g \cap k$. Then c intersects each curve of $G(k)$ beyond Q in at least one direction on k. Hence $\limsup |v| = \infty$ on c. A similar argument shows that $\limsup |v| = \infty$ on c if R has two boundary chains and c intersects each curve of $F[R]$. Hence we conclude that the function $f = u + iv$ is an allowable function contoured by (F, G) which has no finite asymptotic value. The theorem now follows by the remark at the end of Section 1.2.8.

Note. The only property of v and G used in the proof of Theorem 3.6-1 which is not included in the hypothesis of Theorem 3.5-1 is that v differs by ± 1 at the endpoints of a branched curve, and this property is only used when the set of branch points on chains connecting the intersections of the path c and a tree T is unbounded.

As a consequence of Corollary 2.2-2-1 and the observation following it, we obtain the following corollaries.

COROLLARY 3.6-1-1. *F contours the real part of an analytic function with no finite asymptotic value if the trees of F are isolated and F contains no cyclic set V containing three or more elements such that $\theta(V)$ contains no branched curve.*

COROLLARY 3.6-1-2. *F contours the real part of an analytic function with no finite asymptotic value if the trees of F are isolated and each regular curve or maximal chain of F tends properly to infinity relative to some family complementary to F.*

3.7. *Curve families and conjugate nets contouring polynomials.*

LEMMA 3.7-1. *If F has at most a finite number of branch points and contours the real part of an allowable function $w = f(z) = u + iv$ which has no finite asymptotic value, then f is topologically equivalent to a polynomial.*

Proof. By Theorem 1.2-9, f is topologically equivalent to a function g analytic in π or in $|z| < 1$. We define $f(\infty) = \infty$ and shall then show that f is a light interior mapping of the extended plane onto the extended plane. Hence by Stoïlow's theorem [12, p. 121], f is topologically equivalent to an analytic mapping g of the extended plane onto the extended plane. We can leave ∞ fixed under the homeomorphism relating f and g, so that $g(\infty) = \infty$ also, and g has no other pole. Hence g is a polynomial.

We first show f is continuous at ∞. Since f has no finite asymptotic value and u is monotone on any regular curve or maximal chain of G, $|u|$ tends to ∞ in both directions on any element of G and the same holds for v on elements of F. It follows that the locus $|f| \leqq N$ is bounded in any region R. Since there are but a finite number of regions the locus $|f| \leqq N$ is a closed bounded subset of π. Therefore the inverse image of the set in the w-plane $|w| > N$ is an open set containing $z = \infty$. Hence f is continuous at ∞. Also from the monotonicity of $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ on cross-sections of $F[R]$ and $G[R]$ respectively for each region R, it follows that f assumes a given value at most at a finite number of points of π, hence f is light, and that each point of the extended w-plane is covered by f.

The function f is interior at each finite point since it is an allowable function. Fix $r > 0$ and let $\rho = \max |f(z)|$ for $|z| \leqq r$. Then for each w in $|w| > \rho$ there exists a z such that $f(z) = w$ by the preceding paragraph; furthermore $|z| > r$. Hence the image of any neighborhood of $z = \infty$ contains a neighborhood of $w = \infty$, so f is interior at ∞ also.

COROLLARY 3.7-1-1. *If (F, G) is a conjugate net, F has at most a finite number of branch points and (F, G) contours an allowable function $f = u + iv$ which has no finite asymptotic value, then f is topologically equivalent to a polynomial.*

THEOREM 3.7-1. *Necessary and sufficient conditions that a branched regular curve family F contour a harmonic polynomial are that F contain at most a finite number of branch points and no b-curve.*

Proof. The necessity follows from Corollary 2.1-1-1 and the fact that a polynomial has at most a finite number of critical points, hence the corresponding level curve families of the real and imaginary parts of the polynomial have at most a finite number of branch points.

Define u, v and G as in Sections 3.4 and 3.5. Then $f = u + iv$ has no finite asymptotic value and is topologically equivalent to a polynomial g by Lemma 3.7-1. Hence F is homeomorphic to the level curve family of $\mathrm{Re}(g)$ or $\mathrm{Im}(g)$.

THEOREM 3.7-2. *If F contains at most a finite number of branch points, G is complementary to F and each regular curve or infinite chain of G tends properly to infinity relative to F, then (F, G) contours a polynomial.*

Proof. Since each curve G tends properly to infinity relative to F, F and G contain no b-curves as a consequence of the note following Theorem 2.1-1. Hence by Theorem 3.7-1, F is homeomorphic under a homeomorphism h to the level curve family of a harmonic polynomial u. Let v be a harmonic conjugate of u, G' be the family of level curves of v and G'' be the image under h^{-1} of G'. Then G'' is complementary to F, and the conjugate net (F, G'') contours an allowable function $f = u' + iv'$ which is topologically equivalent under h to the polynomial $u + iv$. Hence f has no finite asymptotic value. Suppose c is a regular curve or infinite chain of G. Since F has at most a finite number of branch points, c tends properly to infinity in a region R. Hence $f[R] = F(c \cap R)$. Since (F, G'') contours a function with no finite asymptotic value, each element of G'' tends properly to infinity relative to F, and by Theorem 3.1-2, there is a curve x of G'' such that $F[R] = F(c \cap R) = F(x \cap R)$. Therefore, since $|u'|$ tends to ∞ on $x \cap R$, $|u'|$ also tends to ∞ on $c \cap R$. By interchanging the roles of F and G it follows that there is a function v'' contoured by G which is topologically equivalent to a harmonic polynomial and $|v''|$ tends to ∞ in both directions on any element of F. The families F and G and the functions u' and v'' satisfy the hypotheses of Theorem 3.5-1, hence as a consequence of the note following Theorem 3.6-1, $u' + iv''$ is an allowable function contoured by (F, G) which has no finite asymptotic value. By Lemma 3.7-1, $u' + iv''$ is topologically equivalent to a polynomial.

Note. The hypotheses of Theorem 3.7-2 are also necessary conditions that (F, G) contour a function topologically equivalent to a polynomial. Hence the hypotheses of Theorem 3.7-2 characterize the conjugate nets which contour a polynomial, while the hypotheses of Theorem 3.7-1 characterizes the curve families which contour a harmonic polynomial.

UNIVERSITY OF UTAH.

REFERENCES.

[1] L. Bieberbach, *Lehrbuch der funktionentheorie*, Volume II, Leipzig (1931).

[2] W. Boothby, "The topology of regular curve families with multiple saddle points," *American Journal of Mathematics*, vol. 73 (1951), pp. 405-438.

[3] ———, "The topology of the level curves of harmonic functions with critical points," *American Journal of Mathematics*, vol. 73 (1951), pp. 512-538.

[4] E. Collingwood and M. Cartwright, "Boundary theorems for a function meromorphic in the unit circle," *Acta Mathematica*, vol. 87 (1952), pp. 83-146.

[5] H. Cullen, *A set of parabolic regular curve families filling the plane and certain related Riemann surfaces*, thesis, Ann Arbor (1949).

[6] F. Iversen, *Recherches sur les fonction inverses des fonctions méromorphes*, thesis, Helsinki (1914).

[7] J. Jenkins and M. Morse, "Conjugate nets on an open Riemann surface," *Lectures on functions of a complex variable*, Ann Arbor (1955), pp. 123-185.

[8] W. Kaplan, "Regular curve families filling the plane, I," *Duke Mathematical Journal*, vol. 7 (1940), pp. 154-185.

[9] ———, "Regular curve families filling the plane, II," *Duke Mathematical Journal*, vol. 8 (1941), pp. 11-46.

[10] M. Morse, *Topological Methods in the theory of functions of a complex variable*, Annals of Mathematics Studies, Number 15, Princeton (1947).

[11] R. Nevanlinna, *Eindeutige analytische funktionen*, Second edition, Berlin (1953).

[12] S. Stoïlow, *Leçons sur les principes topologiques de la théorie des fonctions analytiques*, Paris (1938).

[13] D. Wend, *Branched regular curve families and finite asymptotic paths of analytic functions*, thesis, Ann Arbor (1954).

[14] G. T. Whyburn, *Analytic topology*, American Mathematical Society Colloquium Publications, Volume 28, New York (1942).

DIOPHANTINE APPROXIMATIONS ON TORUSES.

By Serge Lang.

The main purpose of this paper is to make precise certain analogies between diophantine approximations on the multiplicative group and elliptic curves, and to show how these are related with approximations on the additive group by means of the exponential mapping. We shall state two conjectures. To relate their exponential and logarithmic formulations, it is best to prove first some results relating the exponential and logarithmic heights.

1. Heights on the multiplicative group. Let K be a number field, and let M_K be the set of absolute values on K extending the standard, normalized absolute values on the rational numbers $\boldsymbol{Q}$. (cf. [3], Chapter III.) If $|\ |_v$ is in M_K, and α is an element of K, we write

$$v(\alpha)=\log\|\alpha\|_v$$

where $\|\alpha\|_v=|\alpha|_v^{N_v}$. Instead of saying that $|\ |_v$ is in M_K, we shall also say that v is in M_K. By the product formula, we have

$$\sum_v v(\alpha)=0$$

for any $\alpha\in K^*$ (= multiplicative group of K). If $(\alpha_0,\cdots,\alpha_n)$ is a point in projective n-space with coordinates in K, we define its *height* to be

$$\sum_v \sup_i(0, v(\alpha_i)).$$

(This is the logarithm of the height as used frequently in the past literature.) If P is the point represented by $(\alpha_0,\cdots,\alpha_n)$, we shall write $h(P)$ for its height (again taking the log of the notation in [3]). If α is an element of K, we let $h(\alpha)$ be the height of the point $(1,\alpha)$ in projective 1-space.

We shall use the standard Vinogradov notation for orders of magnitude. If ϕ, ψ are two real functions on a set, whose values are $\geqq 0$, we write $\phi \ll \psi$ if there exists a constant $c>0$ such that $\phi\leqq c\psi$. We write $\phi \gg\ll \psi$ if $\phi\ll\psi$ and $\psi\ll\phi$.

Let Γ be a finitely generated subgroup of the multiplicative group K^* of K. There exists a finite set $S\supset S_\infty$ of absolute values in M_K, such that Γ

Received November 1, 1963.

is contained in the S-units K_S, i.e. those elements of K^* having absolute value equal to 1 outside S. Let $s+1$ be the cardinality of S, and map

$$L\colon K_S \to \boldsymbol{R}^{s+1}$$

by the log, such that

$$L(\beta) = (v_1(\beta), \cdots, v_{s+1}(\beta)),$$

if v_j are the elements of S. The kernel of L consists precisely of the roots of unity in K. The image of L is a lattice of dimension s in the hyperplane H defined by the equation

$$\sum_{j=1}^{s+1} x_j = 0.$$

We now define a function $\boldsymbol{F}$ on $\boldsymbol{R}^{s+1}$ by the condition

$$\boldsymbol{F}(X) = \sum_{j=1}^{s+1} \sup(0, x_j).$$

Then for any $t \geqq 0$ we have $\boldsymbol{F}(tX) = t\boldsymbol{F}(X)$, and for any two vectors X, Y we have $\boldsymbol{F}(X+Y) \leqq \boldsymbol{F}(X) + \boldsymbol{F}(Y)$. We view $\boldsymbol{F}$ as restricted to the hyperplane H. Then for X in H, the condition $\boldsymbol{F}(X) = 0$ implies $X = 0$. Finally, for X in H, we have $\boldsymbol{F}(-X) = \boldsymbol{F}(X)$. Hence $\boldsymbol{F}$ is a norm function on H, and in particular, induces a norm function on any subspace of H.

The image of Γ under L is a discrete subgroup of $\boldsymbol{R}^{s+1}$. Let W be the vector space generated by this image. Then $L(\Gamma)$ is a lattice of maximal rank, say m, in the vector space W. Let $B_1, \cdots, B_m$ be a basis of $L(\Gamma)$. For any vector X in W, we can write

$$X = y_1 B_1 + \cdots + y_m B_m$$

We define

$$\boldsymbol{G}(X) = \sup |y_i|.$$

Then $\boldsymbol{G}$ is a norm on W. Since two norms on a finite dimensional vector space are always of the same order of magnitude, we conclude:

PROPOSITION 1. *We have $\boldsymbol{F} \gg\ll \boldsymbol{G}$ on the space W.*

Let $\beta_1, \cdots, \beta_m$ be free generators of Γ modulo its torsion group. For each β in Γ we can write

$$\beta = \beta_1^{\nu_1} \cdots \beta_m^{\nu_m} \tau$$

with suitable integers $\nu_1, \cdots, \nu_s$ and a root of unity τ. Define

$$\nu = \nu(\beta) = \sup |\nu_i|.$$

From Proposition 1, we obtain immediately the

COROLLARY. *As functions on* Γ, *we have* $v >\!\!> \!\!< \!\!< h$.

We shall now give an asymptotic formula for the number of units of bounded height. This amounts to counting lattice points in certain regions. It seems to me that for such problems, the criterion given by Davenport in [1] is rather hard to apply, and that it is better to proceed as follows. One proves that the boundary of the region can be parametrized by a finite number of unit cubes whose dimension is smaller than that of the region. For homogeneous problems, one can then use the following result.

PROPOSTION 2. *Let* D *be a bounded set in* $\mathbf{R}^m$ *and assume that its boundary is contained in a finite union of pieces, each of which is the image of the unit* $(m-1)$*-cube under a mapping satisfying a Lipschitz condition. Let* Λ *be a lattice in* $\mathbf{R}^m$ *of dimension* m. *Then for* $t \to \infty$, *the number of lattice points in* tD *is equal to*

$$\frac{\operatorname{Vol} D}{\operatorname{Vol} F} t^m + O(t^{m-1})$$

where Vol *means volume in* $\mathbf{R}^m$, *and* F *is a fundamental domain for the lattice.*

Proof. We recall that the boundary is taken in the sense of point set topology. It is the set of points having arbitrary small neighborhoods intersecting both the set D and its complement.

Let $\lambda_1, \cdots, \lambda_m$ be a basis for Λ over $\mathbf{Z}$, and let F be the canonical fundamental domain consisting of all vectors

$$t_1\lambda_1 + \cdots + t_m\lambda_m, \qquad 0 \leqq t_j < 1.$$

If λ is an element of the lattice, and if the translation F_λ of F intersects tD but is not contained in tD, then F_λ intersects the boundary of tD. To see this, let a be a point in $F_\lambda \cap tD$ and let b be a point in F_λ but not in tD. Since F_λ is convex, the line segment between a and b is contained in F_λ. There exists at least one point on this line segment which intersects the boundary of tD. Indeed, if a is an interior point of tD, we consider the set of points on the segment lying in the interior of tD. It is non-empty and open. If no point of this segment lies in the closure of tD but not in the interior of tD, then our set is also closed, contradiction, since the segment is connected, and b lies outside tD.

Let:

$n(t)=$ number of translates F_λ contained in tD,
$n'(t)=$ number of translates F_λ intersecting tD,
$b(t)=$ number of translates F_λ intersecting $\partial(tD)$.

Then $0\leqq n'(t)-n(t)\leqq b(t)$. Each translate F_λ of F contains precisely one lattice point, namely λ. If we can prove that

$$b(t)=O(t^{m-1}),$$

then the assertion of the proposition follows at once from the inequality

$$\operatorname{Vol}(F)n(t)\leqq\operatorname{Vol}(tD)=t^m\operatorname{Vol}(D)\leqq\operatorname{Vol}(F)n'(t).$$

We now prove that $b(t)\ll t^{m-1}$. Let $g\colon I^{m-1}\to \boldsymbol{R}^m$ be a mapping of the $(m-1)$-cube I^{m-1} satisfying a Lipschitz condition, and parametrizing one piece of the boundary ∂D. Since $\partial(tD)=t\partial(D)$, the map tg parametrizes the corresponding piece of $\partial(tD)$. It suffices to prove our assertion for this piece. Cut up the sides of I^{m-1} into intervals of equal length $1/[t]$. This determines a partition of I^{m-1} into $\ll t^{m-1}$ small cubes, each such small cube having a diameter $\leqq c_1/t$ for some constant c_1. The image of such a small cube under g has diameter $\leqq c_2/t$ (where c_2 is equal to c_1 times a Lipschitz constant for g). Hence the image of such a small cube under tg has diameter $\leqq c_2$. Hence such an image meets only a bounded number of translates F_λ of our fundamental domain. Since we have at most c_3t^{m-1} small cubes, we have proved our assertion.

Note. It would be a useful criterion if one knew that a compact subset of a real algebraic set in $\boldsymbol{R}^m$ of algebraic dimension $<m$ satisfied the hypothesis of Proposition 2. Because of the singularities, it does not seem easy to prove, and is one source of hand-waving in the literature. *Added in proof*: I understand from Hironaka that he can prove this statement.

We shall apply Proposition 2 to count units in number fields.

Let S be a finite subset of M_K containing the archimedean absolute values. Let $s+1$ be the cardinality of S, and let K_S be the S-units, i. e. the element of K^* whose absolute value outside S is equal to 1. Then K_S modulo a finite group of roots of unity is a free abelian group on s generators.

PROPOSITION 3. *Let* Γ *be a finitely generated subgroup of* K_S. *For each* $t>0$, *let* $n_\Gamma(t)$ *be the number of elements of* Γ *whose height is* $\leqq t$. *Then there is a constant* $c>0$ *such that*

$$n_\Gamma(t)=ct^m+O(t^{m-1})$$

where m *is the rank of* Γ.

Proof. As before, we map K_S into Euclidean $(s+1)$-space $\mathbf{R}^{s+1}$ by the log mapping:

$$L\colon K_S \to \mathbf{R}^{s+1}$$

such that

$$L(\beta) = (v_1(\beta), \cdots, v_{s+1}(\beta)),$$

if v_j are the elements of S. The kernel of L consists precisely of the roots of unity in K. The image of L is a lattice of dimension s in the hyperplane H defined by the equation

$$\sum_{j=1}^{s+1} x_j = 0.$$

Let $D(t)$ be the region consisting of points $(x_1, \cdots, x_{s+1})$ contained in the hyperplane, and satisfying the relation

$$\sum_{j=1}^{s+1} \sup(0, x_j) \leqq t.$$

Then $D(t) = tD(1)$, and up to a root of unity, the elements of K_S of height $\leqq t$ can be identified with the lattice points contained in $D(t)$. The set of vectors X in the hyperplane such that no x_j is equal to 0 and

$$\sum_{j=1}^{s+1} \sup(0, x_j) < 1$$

is open in H. Hence the boundary of $D(1)$ consists of the vectors X in $D(1)$ satisfying

$$\sum_{j=1}^{s+1} \sup(0, x_j) = 1,$$

except for pieces where some $x_j = 0$, which can be dealt with by induction. Let T denote a subset of S (identified with the integers $1, \cdots, s+1$) which is neither empty nor equal to S. The points $(x_1, \cdots, x_{s+1})$ satisfying

$$x_\nu \geqq 0, \ \sum_{\nu \in T} x_\nu = 1 \quad \text{and} \quad x_\mu \leqq 0, \ \sum_{\mu \notin T} x_\mu = -1$$

are contained in the boundary, and the boundary is the union of such sets of points as T ranges over subsets of S as above, together with pieces with some $x_j = 0$ which can be dealt with inductively.

Let W be the subspace generated by the elements of Γ over the real numbers. Then Γ is a discrete subgroup of W, of rank m, i. e. Γ is a lattice of maximal rank in W. We also observe that W is a subspace of H. Let $D_W(t) = D(t) \cap W$. The interior (open) of $D(t)$ intersected with W is open in W. Hence the boundary of $D_W(t)$ is contained in the intersection of the boundary of $D(t)$ and W. Since $D_W(t) = tD_W(1)$, we are led to considering the finite union of pieces described by a set T as above, intersected with W, and we must show that each one of these can be parametrized by an $(m-1)$-cube. If we can do this, we can then apply Proposition 1 to conclude our proof.

Without loss of generality, say that T consists of the first p coordinates for some integer p. Let

$$B = (1, \cdots, 1, 0, \cdots, 0) \text{ and } B' = (0, \cdots, 0, 1, \cdots, 1)$$

be the vectors having coordinates equal to 1 in the first p places (resp. the last $s+1-p$). Let $A_0 = B + B'$. Then H is the hyperplane perpendicular to A_0. We can find vectors $A_1, \cdots, A_{s-m}$ in H such that $A_0, A_1, \cdots, A_{s-m}$ are mutually perpendicular, and such that W consists of all vectors of H perpendicular to $A_1, \cdots, A_{s-m}$. Let E be the set of vectors X such that

$$X \cdot B = 1 \text{ and } X \cdot B' = -1.$$

Then $D(1)$ is contained in E, and is obtained from E by imposing additional inequalities, namely $x_\nu \geqq 0$ for $\nu \in T$ and $x_\mu \leqq 0$ for $\mu \notin T$. The intersection $E \cap W$ is the set of vectors X such that:

$$X \cdot B = 1,\ X \cdot B' = -1,\ X \cdot A_1 = 0, \cdots, X \cdot A_{s-m} = 0.$$

We must now consider two cases.

Case 1. The vectors B, B', $A_1, \cdots, A_{s-m}$ are linearly independent. Then E is the translation of an $(m-1)$-dimensional space, and applying the additional inequalities to E yields an $(m-1)$-dimensional closed, bounded intersection of planes and half-planes which can be parametrized by an $(m-1)$-cube. (We leave it to the reader to select some parametrization.)

Case 2. The vectors $B, B', A_1, \cdots, A_{s-m}$ are linearly dependent. Since $A_0, \cdots, A_{s-m}$ are mutually perpendicular, and $A_0 = B + B'$, there exists numbers b, $a_1, \cdots, a_{s-m}$ such that

$$B' = bB + a_1 A_1 + \cdots + a_{s-m} A_{s-m}.$$

Let X be in $E \cap W$. Taking the scalar product with X, we obtain

$$X \cdot B' = -1 = b.$$

Hence $B + B' = A_0$ is a linear combination of $A_1, \cdots, A_{s-m}$, which is impossible. We conclude that in Case 2, the intersection $E \cap W$ is empty, and the intersection of the boundary of $D(1)$ with W corresponding to the piece relative to T is empty.

In both cases, we see that $\partial D(1) \cap W$ admits a parametrization by an $(m-1)$-cube. Thus Proposition 1 is applicable and Proposition 2 is proved.

(*Note.* The analogous counting theorem for elements of K of bounded height is considerably harder to prove, and has been done by S. Schanuel in [4].)

2. Approximation on the multiplicative group. Let G be the multiplicative group. Let Γ be a finitely generated subgroup of G_K, where K is a number field. Let ϕ be a non-constant rational function on G, defined over K, and let r be the maximum of the multiplicities of its zeros on G (so distinct from 0 or ∞). Let m be the rank of Γ.

CONJECTURE 1. *The height of points P in Γ which are bounded away from* 0 *and* ∞ *and satisfy the inequality*

$$|\phi(P)| \ll \frac{1}{h(P)^{rm+\epsilon}} \qquad (\epsilon > 0) \tag{1}$$

is bounded.

We shall transform inequality (1). The function field $K(G)$ is generated by a single function, and ϕ can be viewed as a rational function. A point P is represented by an element of K, say β, and if $\phi(\beta)$ is small in absolute value, then there is some root $\alpha \neq 0$ of the rational function ϕ such that β is close to α. If β is close to α, then its distance from any other zero or pole of ϕ is bounded away from 0 (approximately by the distance of α itself from another zero or pole). The multiplicity of α in a factorization of ϕ is at most r. The worst possible case is that in which this multiplicity is r. In that case, $|\phi(\beta)|$ is approximately equal to $|\alpha-\beta|^r$ (up to a constant factor). Hence our inequality amounts to

$$|\alpha-\beta| \ll \frac{1}{h(\beta)^{m+\epsilon}}.$$

If one replaces $h(\beta)^{m+\epsilon}$ by $e^{h(\beta)\epsilon}$ in the right hand side of our inequality, then the statement is actually a known theorem of Mahler (cf. for instance [3], Corollary of Theorem 1, Chapter VI, § 1).

The above approximation can be transferred to the additive group. Let $\beta_1, \cdots, \beta_m$ be free generators of Γ modulo its torsion group (i. e. the group of roots of unity). Then we can write

$$\beta = \beta_1^{\nu_1} \cdots \beta_m^{\nu_m} \tau$$

for some torsion element τ. Furthermore, $|\alpha - \beta|$ is of the same order of magnitude as $|\log \alpha - \log \beta|$. Since τ ranges over a finite set, in dealing with solutions of the inequality (1), we may assume that we have always the same τ. Let $u_0 = \log(\alpha\tau)$. Taking the corollary of Proposition 2 into account, we may rewrite our inequality in the form

$$|u_0 - \nu_1 \log \beta_1 - \cdots - \nu_m \log \beta_m + \nu_{m+1} 2\pi i| << \frac{1}{\nu^{m+\epsilon}},$$

where the log is one fixed value of the logarithm.

We have therefore transferred our diophantine approximation on the multiplicative group over a number field into an inhomogeneous approximation on the additive group. The period $2\pi i$ of the exponential function contributes one term to the sum on the left, and gives rise to $m+1$ free choices of the coefficients $\nu_1, \cdots, \nu_{m+1}$. A standard application of Dirichlet's principle shows that m cannot be replaced by a smaller exponent on the right hand side. In fact, given real numbers $\xi_1, \cdots, \xi_m$ and an integer $\nu > 0$, there exist integers $\nu_1, \cdots, \nu_m$ not all 0, such that

$$|\nu_1 \xi_1 + \cdots + \nu_m \xi_m| << \frac{1}{\nu^{m-1}},$$

and $|\nu_j| \leqq \nu$. (This has content of course only when $\xi_1, \cdots, \xi_m$ are linearly independent over the rationals.) Hence the exponent rm in the conjecture cannot be improved upon.

Khintchine's theorem and suitable generalizations by Schmidt ([5] Theorem 2) imply the following: For almost all sets $\xi_1, \cdots, \xi_m, \eta$ of real numbers the solutions of the inequality

$$|\eta + \nu_1 \xi_1 + \cdots + \nu_m \xi_m| << \frac{1}{\nu^{m-1+\epsilon}}$$

(where $\nu = \max |\nu_j|$) are finite in number. Thus our conjecture essentially asserts that the logarithms of algebraic numbers behave like "almost all numbers". It may be viewed as expressing a measure of linear independence for logarithms of algebraic numbers. One conjectures of course that $\log \beta_1, \cdots, \log \beta_m$ are in fact algebraically independent, and one could make an analogous conjecture for their measure of algebraic independence.

Actually, in view of the fact that the period of the exponential is pure imaginary, namely $2\pi i$, the approximation in its additive form separates into two simultaneous approximations, concerning the real part and the pure imaginary. If u_0, $\log \beta_1, \cdots, \log \beta_m$ are all real, then to get a small number on the left hand side of our inequality, we must have $\nu_{m+1} = 0$. In that extreme case, the exponent m will be wasteful, and should be replaced by $m - 1$.

Gelfond [2] has obtained transcendence measures for quotients of logarithms of algebraic numbers which improve the Mahler result in that case, for instance of the type

$$| \nu_1 \log \alpha + \nu_2 \log \beta | \gg \frac{1}{\nu^{(\log \nu)^{2+\epsilon}}}.$$

for algebraic α, β multiplicatively independent. This is still far short of what is to be expected, but Gelfond's clever method might be a step in the right direction. In any case, it is a first indication that the Mahler result is a priori too coarse.

I have had some preliminary computations comparing the absolute values of expressions of type

$$| \nu_1 \log 2 + \nu_2 \log 3 + \nu_3 \log 5 |$$

with $1/\nu^{2+\epsilon}$ (letting $\nu = \max | \nu_j |$), and $\epsilon = 1/4$. They were carried out only for $\nu \leqq 50$. They seem to indicate a tendency for the solutions to thin out, but one should carry them much further to get a significant computational support for the conjecture that there is only a finite number.

Actually, the computations suggest one other possibility. For each ν, let $\psi(\nu)$ be the number of solutions of the inequality

$$| \nu_1 \log \beta_1 + \cdots + \nu_n \log \beta_n | \leqq \frac{1}{\nu^{n-1}}, \qquad | \nu_i | \leqq \nu$$

where $\beta_1, \cdots, \beta_n$ are, say, real algebraic numbers. Then in the above case, except for $\nu = 1$, the values of $\psi(\nu)$ were 1, 2, or 3. Thus there is a vague possibility that $\psi(\nu)$ might be bounded. It would also be interesting to push the computations quite far, to see how $\psi(\nu)$ behaves asymptotically. Especially for this function, ν ranging up to 50 cannot be regarded as really significant.

3. Heights on elliptic curves.

Aside from the additive and multiplicative groups, the only other group variety of dimension 1 is an elliptic curve. We shall be interested to the analogues of Proposition 3 and Conjecture 1 for such curves. In this section we deal with the analogue of Proposition 3, and I shall reproduce results of Tate (unpublished at the moment, as usual). It will in fact be just as easy to work on abelian varieties.

Two functions on a set will be called *equivalent* if their difference is bounded in absolute value.

THEOREM 1. (Tate) *Let A be an abelian variety in projective space, defined over the number field K. Then there exists a unique positive definite quadratic form q and a linear form l on A_K such that h is equivalent to $q+l$, as functions on A_K.*

Proof. We consider the function f on $A \times A$ given by $f(P,Q) = h(P+Q) - h(P) - h(Q)$. We contend that as a function of each variable, f is essentially linear, i. e. for all P_1, P_2, Q the expression

$$|f(P_1+P_2,Q) - f(P_1,Q) - f(P_2 Q)|$$

is bounded. This expression is equal to

$$\begin{aligned} |h(P_1+P_2+Q) - h(P_1+P_2) - h(P_1+Q) \\ - h(P_2+Q) + h(P_1) + h(P_2) + h(Q)|. \end{aligned}$$

Let $\phi_{123}\colon A \times A \times A \to A$ be the sum. Let ϕ_{12} be the composite map

$$\begin{tikzcd} A \times A \times A \arrow[r, "\pi_{12}"] & A \times A \arrow[r, "s_2"] & A \end{tikzcd}$$

where π_{12} is the projection on the first two factors and s_2 is the sum, and define ϕ_{13}, ϕ_{23} similarly. Let ϕ_i $(i = 1, 2, 3)$ be the projection on the i-th factor. Let H be a hyperplane section giving the projective embedding of A and the height. Then by the theorem of the square,

$$\begin{aligned} \phi_{123}^{-1}(H) - \phi_{12}^{-1}(H) - \phi_{13}^{-1}(H) - \phi_{23}^{-1}(H) \\ + \phi_1^{-1}(H) + \phi_2^{-1}(H) + \phi_3^{-1}(H) \end{aligned}$$

is linearly equivalent to 0. It is a divisor on $A \times A \times A$, and the height associated with this divisor is therefore equivalent to 0, by one of the main properties of heights (cf. [3], Property 3, Chapter IV, § 2). This height is precisely the function expressing the linearity of f in the first variable.

We now define

$$b(P,Q) = \lim_{n\to\infty} \frac{1}{2^{2n}} f(2^n P, 2^n Q).$$

It is a trivial exercise to verify that $b(P,Q)$ is actually bilinear, and is equivalent to $f(P,Q)$.

We let $q(P) = \frac{1}{2} b(P,P)$. It is a trivial exercise to verify that $h - q$ is equivalent to a unique linear functional l. Thus it follows that h is equivalent to $q + l$. Since the height is always $\geqq 0$ and there is only a finite number of points with bounded height, it follows that q is positive definite. The uniqueness statements are trivially proved.

COROLLARY 1. (Tate) *If the hyperplane section H of A giving the projective embedding satisfies $H = H^-$ (transform of H under the automorphism $P \rightsquigarrow -P$ of A), then the linear function l is equal to* 0.

Proof. Easy exercise, using the theorem of the square.

COROLLARY 2. *Let Γ be a finitely generated subgroup of A_K. Let $P_1, \cdots, P_m$ be free generators of Γ modulo the torsion group, and for each P in Γ, write*

$$P = \nu_1 P_1 + \cdots + \nu_m P_m + Q$$

where Q is an element of the torsion group. Let $\nu = \nu(P) = \max |\nu_i|$. As function on Γ, we have $h >\!\!><\!\!< \nu^2$.

Proof. Obvious.

COROLLARY 3 (Néron). *Let Γ be as in Corollary 2. The number of points P in Γ such that $q(P) \leqq t$ is equal to*

$$ct^{m/2} + O(t^{(m-1)/2}) \qquad \textit{for } t \to \infty$$

and some constant $c > 0$.

Proof. In the case of a positive definite quadratic form q on Γ, we can view q as defined on the tensor product of Γ with the real numbers, and the condition of parametrizability stated in Proposition 2 is satisfied in the present situation, because the unit sphere relative to the quadratic form is a compact submanifold of dimension $m - 1$ in m-space.

4. Approximation on elliptic curves. Let A be an abelian variety of dimension 1 defined over the number field K. We can represent A (or rather its complex points) as a quotient of the complex 1-space, say by means of a $\wp$-function associated with A, with g_2, g_3 in K. If u is a complex number, we denote by $\exp(u)$ the point whose projective coordinates are $(1, \wp(u), \wp'(u))$. Conversely, if P is a point on A, we let $\log P$ be some fixed u such that $\exp(u) = P$ (we choose u in a fixed, bounded fundamental domain).

Let ϕ be a non-constant rational function on A, defined over the number field K and let r be the maximum of the multiplicities of its zeros. Le Γ be a finitely generated subgroup of A_K, of rank m.

CONJECTURE 2. *The height of points P in Γ which satisfy the inequality*

$$|\phi(P)| << \frac{1}{h(P)^{\frac{1}{2}r(m+1)+\epsilon}}$$

is bounded.

If one replaces the right hand side of the inequality by $e^{h(P)\epsilon}$, then it is known that there is only a finite number of P satisfying the corresponding inequality. This is essentially the formalization of one step in the Siegel argument that an elliptic curve has only a finite number of integral points, taking Roth's Theorem into account (cf. Theorem 1 of [3], Chapter VII, § 1).

The rational function ϕ may be viewed as a doubly periodic elliptic function, so that we have $\phi(P) = \Phi(u)$, for u complex. Since the complex points of A form a compact set, any infinite sequence of points P whose height tends to infinity and satisfies our inequality must have a point of accumulation P which will be a zero of ϕ. The worst case from our point of view is when Φ has a zero of order r at P_0. In that case, Φ has a power series expansion near $u_0 = \log P_0$ given by

$$\Phi(u) = c_r(u - u_0)^r + \cdots$$

for some constant $c_r \neq 0$.

Let $P_1, \cdots, P_m$ be free generators of A_K modulo the torsion group. Let $u_j = \log P_j$. Let ω_1, ω_2 be two fundamental periods. Let

$$P = \nu_1 P_1 + \cdots + \nu_m P_m + Q$$

where Q is in the torsion group. Without loss of generality, we may assume that we always have the same point Q, for P satisfying our inequality. Then P is close to P_0, and there exist integers ν_{m+1}, ν_{m+2} such that

$$\nu_1 u_1 + \cdots + \nu_m u_m + \nu_{m+1}\omega_1 + \nu_{m+2}\omega_2$$

is close to $u_0 = \log(P_0 - Q)$. Hence our inequality amounts to

$$|-u_0 + \nu_1 u_1 + \cdots + \nu_m u_m + \nu_{m+1}\omega_1 + \nu_{m+2}\omega_2| << \frac{1}{h(P)^{\frac{1}{2}(m+1)+\epsilon}}.$$

This is an inhomogeneous approximation problem near u_0. By Tate's theorem, h and ν^2 are of the same order of magnitude (letting ν as usual be $\max|\nu_j|$). Hence our inequality may be rewritten as

$$|-u_0 + \nu_1 u_1 + \cdots + \nu_m u_m + \nu_{m+1}\omega_1 + \nu_{m+2}\omega_2| << \frac{1}{\nu^{m+1+\epsilon}}.$$

Thus our conjecture asserts that from the point of view of diophantine approximations, the logs of algebraic points on our elliptice curve behave again like "almost all numbers".

Actually, the two periods ω_1 and ω_2 are linearly independent over the reals. This would seem to indicate that the exponent $m+1$ should be replaced by m, just as for the multiplicative group. This also raises the following question. Given a non-zero period ω, are there infinitely many real values of t such that $\wp(t\omega)$ are algebraic and linearly independent over the integers? The extent to which m can be further lowered depends on the existence of such values of t.

REFERENCES.

[1] H. Davenport, "On a principle of Lipschitz", *Journal of the London Mathematical Society*, vol. 26 (1951), pp. 179-183.

[2] A. O. Gelfond, *Transcendental and algebraic numbers*, translated by Leo F Boron, Dover, 1960.

[3] S. Lang, *Diophantine Geometry*, Interscience, New York, 1962.

[4] S. Schanuel, "Heights in number fields", *to appear.*

[5] W. Schmidt, "A metrical theorem in diophantine approximation", *Canadian Journal of Mathematics*, vol. 22 (1960), pp. 619-631.

INVARIANT DIFFERENTIAL OPERATORS AND DISTRIBUTIONS ON A SEMISIMPLE LIE ALGEBRA.

By HARISH-CHANDRA.

1. Introduction.

1. Introduction. This paper deals with the second main step in the proof of Lemma 1 of [2(d)]. As indicated in [2(d), p. 121] this proof proceeds by induction on $\dim \mathfrak{g}$. Let H be a semisimple element of $\mathfrak{g}$ and $\mathfrak{z}$ the centralizer of H in $\mathfrak{g}$. Then $\dim \mathfrak{z} < \dim \mathfrak{g}$ if H is not in the center of $\mathfrak{g}$. In this paper we give a method of reducing the problem from $\mathfrak{g}$ to $\mathfrak{z}$, locally around H. Since the induction hypothesis is applicable to $\mathfrak{z}$, this enables us to resolve the question around H. Even when $\mathfrak{g}$ is semisimple, $\mathfrak{z}$ is only reductive and therefore, in order to permit induction, we are forced to start from the assumption that $\mathfrak{g}$ is reductive rather than semisimple.

The basic tools for effecting the above-mentioned reduction are contained in Lemma 17, its corollaries and Theorem 2. The proof of Theorem 3 provides a good illustration of how our method actually works in practice.

This work was supported, in part, by a grant from the Sloan Foundation and a contract with the U. S. Army.

2. Some algebraic results. Let k be a field of characteristic zero and $\mathfrak{g}$ a Lie algebra over k. We use the notation of [2(e), §§ 6-7]. If $n = \dim \mathfrak{g}$, we can select $p_i \in P(\mathfrak{g})$ $(0 \leq i \leq n)$ such that

$$\det(t - \operatorname{ad} X) = \sum_{0 \leq i \leq n} p_i(X) t^i \qquad (X \in \mathfrak{g})$$

for an indeterminate t. Note that $p_n = 1$. Let l be the least integer such that $p_l \neq 0$. Then l is called the rank of $\mathfrak{g}$ and an element $X \in \mathfrak{g}$ is called regular (in $\mathfrak{g}$) if $p_l(X) \neq 0$. Similarly X is singular if $p_l(X) = 0$.

Now assume that $\mathfrak{g}$ is reductive and $\mathfrak{z}$ is a subalgebra of $\mathfrak{g}$ such that 1) $\mathfrak{z}$ is reductive in $\mathfrak{g}$ and 2) $\operatorname{rank} \mathfrak{z} = \operatorname{rank} \mathfrak{g}$. Put

$$\zeta(Z) = \det(\operatorname{ad} Z)_{\mathfrak{g}/\mathfrak{z}} \qquad (Z \in \mathfrak{z}),$$

where $(\operatorname{ad} Z)_{\mathfrak{g}/\mathfrak{z}}$ denotes the endomorphism of the factor space $\mathfrak{g}/\mathfrak{z}$ defined by $\operatorname{ad} Z$. Let $\mathfrak{z}'$ be the set of all $Z \in \mathfrak{z}$ where $\zeta(Z) \neq 0$. Since $\mathfrak{z}$ is reductive in $\mathfrak{g}$, we can select a linear subspace $\mathfrak{q}$ of $\mathfrak{g}$ such that 1) $\mathfrak{g} = \mathfrak{z} + \mathfrak{q}$, the sum being direct and 2) $[\mathfrak{z}, \mathfrak{q}] \subset \mathfrak{q}$.

Received November 14, 1963.

Let $\mathfrak{c}$ be the center and $\mathfrak{g}_1$ the derived algebra of $\mathfrak{g}$. Then $\mathfrak{g}_1$ is semisimple and $\mathfrak{g}$ is the direct sum of $\mathfrak{g}_1$ and $\mathfrak{c}$.

LEMMA 1. *$\mathfrak{q}$ is the set of all $Y \in \mathfrak{g}_1$ such that $\operatorname{tr}(\operatorname{ad} Z \operatorname{ad} Y) = 0$ for every $Z \in \mathfrak{z}$. Moreover $\zeta \neq 0$.*

For the proof of this lemma, it is clearly permissible to extend the ground field k to its algebraic closure. Hence we may assume that k is algebraically closed. Fix a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{z}$. Since $\operatorname{rank} \mathfrak{z} = \operatorname{rank} \mathfrak{g}$, $\mathfrak{h}$ is also a Cartan subalgebra of $\mathfrak{g}$. Let Q be the set of all roots of $(\mathfrak{g}, \mathfrak{h})$. For any $\alpha \in Q$, select an element $X_\alpha \neq 0$ in $\mathfrak{g}$ such that $[H, X_\alpha] = \alpha(H) X_\alpha$ for all $H \in \mathfrak{g}$. Then since $[\mathfrak{h}, \mathfrak{z}] \subset \mathfrak{z}$ and $[\mathfrak{h}, \mathfrak{q}] \subset \mathfrak{q}$, it is obvious that

$$\mathfrak{z} = \mathfrak{h} + \sum_{\alpha \in Q_1} k X_\alpha, \qquad \mathfrak{q} = \sum_{\alpha \in Q_2} k X_\alpha$$

where Q_1 is the set of those $\alpha \in Q$ for which $X_\alpha \in \mathfrak{z}$ and Q_2 the complement of Q_1 in Q. Since $\mathfrak{z}$ is reductive, it is obvious that Q_1, and therefore also Q_2, are stable under the mapping $\alpha \to -\alpha$ of roots. Since $X_\alpha \in \mathfrak{g}_1$ $(\alpha \in Q)$ and since X_α is orthogonal to

$$\mathfrak{h} + \sum_{\beta \in Q, \beta \neq -\alpha} k X_\beta$$

under the Killing form, while $\operatorname{tr}(\operatorname{ad} X_\alpha \operatorname{ad} X_{-\alpha}) \neq 0$, the first assertion of the lemma is now obvious. Furthermore

$$\zeta(H) = \prod_{\alpha \in Q_2} \alpha(H) \neq 0$$

if H is a regular element of $\mathfrak{h}$. This proves that $\zeta \neq 0$.

Put $\mathfrak{Q} = \mathfrak{S}(\mathfrak{q})$, $\mathfrak{Q}_d = \mathfrak{S}_d(\mathfrak{q})$ $(d \geqq 0)$ and $\mathfrak{Q}_+ = \mathfrak{S}_+(\mathfrak{q})$.

LEMMA 2. *For any $Z \in \mathfrak{z}'$, Γ_Z defines a bijective mapping of $\mathfrak{Q} \otimes S(\mathfrak{z})$ onto $S(\mathfrak{g})$.*

COROLLARY 1. *Fix $p \in S(\mathfrak{g})$. Then for any $Z \in \mathfrak{z}'$, there exist unique elements $\alpha_Z(p) \in S(\mathfrak{z})$ and $\beta_Z(p) \in \mathfrak{Q}_+ \otimes S(\mathfrak{z})$ such that*

$$p - \alpha_Z(p) = \Gamma_Z(\beta_Z(p)).$$

Moreover[1] $d^0 \alpha_Z(p) \leqq d^0 p$.

COROLLARY 2. *For any $p \in S(\mathfrak{g})$ we can choose an integer $r \geqq 0$ such that the mappings $Z \to \zeta(Z)^r \alpha_Z(p)$ and $Z \to \zeta(Z)^r \beta_Z(p)$ $(Z \in \mathfrak{z}')$ can be extended to polynomial mappings of $\mathfrak{z}$ into $S(\mathfrak{z})$ and $\mathfrak{Q}_+ \otimes S(\mathfrak{z})$ respectively.*

---

[1] $d^0 q$ denotes the degree of q.

The proof of all this is substantially the same as in [2(a), § 7].

Remark. Assume $\mathfrak{z}$ is abelian. Then it is a Cartan subalgeba of $\mathfrak{g}$ and so we denote it by $\mathfrak{h}$ in this case. Since $\mathfrak{g} = \mathfrak{h} + \mathfrak{q}$, it is clear that $\mathfrak{G} = \mathfrak{S}(\mathfrak{q})\mathfrak{S}(\mathfrak{h})$. Moreover it is clear from its definition (see [2(e), § 7]) that

$$\sigma_H(H_0)p - d_H p = 0$$

for $H, H_0 \in \mathfrak{h}$ and $p \in S(\mathfrak{h})$. Hence $\Gamma_H(\mathfrak{Q}\mathfrak{S}_+(\mathfrak{h}) \otimes S(\mathfrak{h})) = \{0\}$ for $H \in \mathfrak{h}$. Therefore since

$$\mathfrak{G}_+ = \mathfrak{S}_+(\mathfrak{g}) = \mathfrak{Q}\mathfrak{S}_+(\mathfrak{h}) + \mathfrak{Q}_+,$$

we conclude that

$$\Gamma_H(\mathfrak{G}_+ \otimes S(\mathfrak{h})) = \Gamma_H(\mathfrak{Q}_+ \otimes S(\mathfrak{h})) \qquad (H \in \mathfrak{h}).$$

LEMMA 3. *Let τ be an automorphism of $\mathfrak{g}$ such that $\mathfrak{z}^\tau = \mathfrak{z}$. Then $\mathfrak{z}'$ is invariant under τ and*

$$\alpha_{Z^\tau}(p^\tau) = (\alpha_Z(p))^\tau, \qquad \beta_{Z^\tau}(p^\tau) = (\beta_Z(p))^\tau$$

for $Z \in \mathfrak{z}'$ and $p \in S(\mathfrak{g})$.

Since τ is an automorphism of $\mathfrak{g}$ which leaves $\mathfrak{z}$ invariant, it is clear that $\zeta(Z^\tau) = \zeta(Z)$ $(Z \in \mathfrak{z})$ and τ maps $\mathfrak{q}$ into itself. Hence $\mathfrak{z}'$ is invariant under τ. We can extend τ uniquely to an automorphism of $\mathfrak{G}$ and then to a bijective linear mapping of $\mathfrak{G} \otimes S(\mathfrak{g})$ by setting $(g \otimes q)^\tau = g^\tau \otimes q^\tau$ $(g \in \mathfrak{G}, q \in S(\mathfrak{g}))$. Then it is obvious that

$$(\Gamma_X(u))^\tau = \Gamma_{X^\tau}(u^\tau)$$

for $X \in \mathfrak{g}$ and $u \in \mathfrak{G} \otimes S(\mathfrak{g})$. From this the second statement of the lemma follows immediately.

3. Some elementary facts about differential operators.

Let E be a vector space over $\boldsymbol{C}$ of finite dimension and U a nonempty open subset of E. Let $\mathfrak{H}(U)$ denote the ring of all holomorphic functions on U. For any $X \in E$, let $\partial(X)$ denote, as usual, the derivation of $\mathfrak{H}(U)$ over $\boldsymbol{C}$ defined as follows. If $f \in \mathfrak{H}(U)$, $\partial(X)f$ is the function

$$Y \to \{df(Y + tX)/dt\}_{t=0} \qquad (Y \in U, t \in \boldsymbol{C}).$$

For any $g \in \mathfrak{H}(U)$, we denote the endomorphism $f \to gf$ of $\mathfrak{H}(U)$ again by g. Let $\mathfrak{D}_h(U)$ denote the algebra of all endomorphisms of $\mathfrak{H}(U)$ over $\boldsymbol{C}$ generated by $\partial(X)$ and g $(X \in E, g \in \mathfrak{H}(U))$. We call $\mathfrak{D}_h(U)$ the algebra of holomorphic differential operators on U. For $f \in \mathfrak{H}(U)$, $D \in \mathfrak{D}_h(U)$ and $X \in U$, we denote, as usual, by $f(X;D)$ the value of Df and X.

Now assume U is connected. Then $\mathfrak{H}(U)$ is an integral domain. Let $\mathfrak{M}(U)$ be the quotient field of $\mathfrak{H}(U)$. For any $X \in E$, $\partial(X)$ can be extended uniquely to a derivation of $\mathfrak{M}(U)$ over $\boldsymbol{C}$. Moreover corresponding to any $g \in \mathfrak{M}(U)$, we have the endomorphism $f \rightarrow gf$ $(f \in \mathfrak{M}(U))$ of $\mathfrak{M}(U)$ over $\boldsymbol{C}$, which we again denote by g. Let $\mathfrak{D}_m(U)$ be the algebra of all endomorphisms of $\mathfrak{M}(U)$ generated by $\partial(X)$ and $g(X \in E, g \in \mathfrak{M}(U))$. Then $\mathfrak{D}_m(U)$ is called the algebra of meromorphic differential operators on U. Fix an element $D \in \mathfrak{D}_m(U)$. It is easy to see that if $Df = 0$ for all $f \in \mathfrak{H}(U)$, then $D = 0$. From this it follows without difficulty that there exists an injective homomorphism $D \rightarrow D_m$ of $\mathfrak{D}_h(U)$ into $\mathfrak{D}_m(U)$ such that $D_m f = Df$ for $D \in \mathfrak{D}_h(U)$ and $f \in \mathfrak{H}(U)$. We may therefore identify $\mathfrak{D}_h(U)$ with a subalgebra of $\mathfrak{D}_m(U)$ under this homomorphism. In particular, we have the algebras $\mathfrak{D}_h(E) \subset \mathfrak{D}_m(E)$.

Let E' be the space dual to E. Then E' and $\partial(E)$ are both contained in $\mathfrak{D}_h(E)$. We denote by $\mathfrak{D}(E)$ the subalgebra of $\mathfrak{D}_h(E)$ generated by 1 and $E' + \partial(E)$. $\mathfrak{D}(E)$ is called the algebra of polynomial differential operators on E. The linear mapping $X \rightarrow \partial(X)$ $(X \in E)$ can be extended uniquely to a homomorphism $\partial\colon S(E) \rightarrow \mathfrak{D}(E)$.

For any $p \in P(E)$ and $q \in S(E)$, let $\langle p, q\rangle$ denote the value of $\partial(q)p$ at the origin so that $\langle p, q\rangle = p(0;\partial(q))$.

LEMMA 4. *The bilinear form* $(p, q) \rightarrow \langle p, q\rangle$ *on* $P(E) \times S(E)$ *is nondegenerate.*

Select a base $(X_1, \cdots, X_n)$ for E and let $(\lambda_1, \cdots, \lambda_n)$ be the dual base for E'. Then it is easy to verify that

$$\langle \lambda_1^{m_1}\lambda_2^{m_2} \cdots \lambda_n^{m_n}, X_1^{m_1'} \cdots X_n^{m_n'} \rangle = 0$$

unless $m_i = m_i'$ $(1 \leqq i \leqq n)$ and

$$\langle \lambda_1^{m_1} \cdots \lambda_n^{m_n}, X_1^{m_1} \cdots X_n^{m_n} \rangle = m_1! m_2! \cdots m_n!.$$

From this the assertion of the lemma is obvious.

COROLLARY 1. *The homomorphism* $\partial\colon S(E) \rightarrow \mathfrak{D}(E)$ *is injective.*

This is an immediate consequence of the lemma.

COROLLARY 2. *Fix* $X_0 \in E$ *and let* q *be an element of* $S(E)$ *such that* $p(X_0; \partial(q)) = 0$ *for all* $p \in P(E)$. *Then* $q = 0$.

Define $p_{X_0}(X) = p(X_0 + X)$ for $X \in E$ and $p \in P(E)$. Then $p \to p_{X_0}$ is an automorphism of $P(E)$ and $\langle p_{X_0}, q\rangle = p(X_0; \partial(q)) = 0 \quad (p \in P(E))$. Hence $q = 0$ from Lemma 4.

It is clear that

$$\partial(X) \circ \lambda - \lambda \circ \partial(X) = \lambda(X) \qquad (X \in E, \lambda \in E').$$

Therefore it follows that $\mathfrak{D}(E)$ is spanned by elements of the form $p\partial(q)$ $(p \in P(E),\ q \in S(E))$. Fix $D \in \mathfrak{D}(E)$ and $X_0 \in E$. Then it follows from Corollary 2 above, that there exists exactly one element $D_{X_0} \in \partial(S(E))$ such that $p(X_0; D) = p(X_0; D_{X_0})$ for all $p \in P(E)$. As usual, we call D_{X_0} the *local expression* of D at X_0.

COROLLARY 3. *Suppose* D *is an element in* $\mathfrak{D}(E)$ *such that* $Dp = 0$ *for every* $p \in P(E)$. *Then* $D = 0$.

Fix $X \in E$. Then $p(X; D_X) = p(X; D) = 0$ for every $p \in S(E)$. Hence $D_X = 0$. Now we can write D in the form $D = \sum_{1 \leqq i \leqq r} p_i \partial(q_i)$ where $p_i \in P(E)$, $q_i \in S(E)$ and $q_1, \cdots, q_r$ are linearly independent over $\boldsymbol{C}$. Then

$$D_X = \sum_i p_i(X) \partial(q_i) = 0$$

and so it follows from Corollary 1 above that $p_i(X) = 0 \quad (1 \leqq i \leqq r)$. This being true for every $X \in E$, we conclude that $p_i = 0 \quad (1 \leqq i \leqq r)$ and therefore $D = 0$.

COROLLARY 4. *Let* γ *denote the linear mapping of* $P(E) \otimes S(E)$ *into* $\mathfrak{D}(E)$ *such that* $\gamma(p \otimes q) = p\partial(q) \quad (p \in P(E),\ q \in S(E))$. *Then* γ *is bijective.*

It is clear from what we have seen above that γ is surjective. On the other hand suppose $D = \sum_{1 \leqq i \leqq r} \gamma(p_i \otimes q_i) \quad (p_i \in P(E),\ q_i \in S(E))$, where $q_1, \cdots, q_r$ are linearly independent over $\boldsymbol{C}$. Then if $D = 0$, we have seen above (during the proof of Corollary 3) that $p_1 = p_2 = \cdots = p_r = 0$. This proves that γ is injective.

We know that E' and $\partial(E)$ are both contained in $\mathfrak{D}(E)$. It follows from Corollary 4 of Lemma 4 that $E' \cap \partial(E) = \{0\}$. Let $\mathfrak{T}(E)$ denote the tensor algebra over $E' + \partial(E)$ and ϕ_E the homomorphisms of $\mathfrak{T}(E)$ into $\mathfrak{D}(E)$ which corresponds to the inclusion mapping of $E' + \partial(E)$ into $\mathfrak{D}(E)$. Let $\mathfrak{U}(E)$ be the ideal in $\mathfrak{T}(E)$ generated by elements of one of the following three types.

$$\begin{array}{llr} (1) & \lambda \otimes \partial(X) - \partial(X) \otimes \lambda + \lambda(X) & (\lambda \in E', X \in E), \\ (2) & \partial(X) \otimes \partial(Y) - \partial(Y) \otimes \partial(X) & (X, Y \in E), \\ (3) & \lambda \otimes \mu - \mu \otimes \lambda & (\lambda, \mu \in E'). \end{array}$$

Then clearly $\mathfrak{U}(E)$ lies in the kernel of ϕ_B.

LEMMA 5. *$\mathfrak{U}(E)$ is exactly the kernel of ϕ_B and therefore*

$$\mathfrak{D}(E) \cong \mathfrak{T}(E)/\mathfrak{U}(E).$$

Select a base $(X_1, \cdots, X_n)$ for E and the dual base $(\lambda_1, \cdots, \lambda_n)$ for E'. Define two injective linear mappings α and β of $P(E)$ and $\partial(S(E))$ respectively into $\mathfrak{T}(E)$ as follows.

$$\alpha(1) = 1, \alpha(\lambda_i^m) = \alpha(\lambda_i^{m-1}) \otimes \lambda_i \qquad (1 \leqq i \leqq n, m \geqq 1)$$

$$\alpha(\lambda_1^{m_1}\lambda_2^{m_2} \cdots \lambda_n^{m_n}) = \alpha(\lambda_1^{m_1}) \otimes \alpha(\lambda_2^{m_2}) \otimes \cdots \otimes \alpha(\lambda_n^{m_n})$$

$$\beta(1) = 1, \beta(\partial(X_i^m)) = \beta(\partial(X_i^{m-1})) \otimes \partial(X_i) \qquad (1 \leqq i \leqq n, m \geqq 1)$$

$$\beta(\partial(X_1^{m_1}X_2^{m_2} \cdots X_n^{m_n})) = \beta(\partial(X_1^{m_1})) \otimes \beta(\partial(X_2^{m_2})) \otimes \cdots \otimes \beta(\partial(X_n^{m_n})).$$

It is obvious that $\phi_B(\alpha(p)) = p$, $\phi_B(\beta(\partial(q))) = \partial(q)$ for $p \in P(E)$ and $q \in S(E)$. Now fix $\xi \in \mathfrak{T}(E)$. It is clear from the definition of $\mathfrak{U}(E)$ that we can choose elements $p_i \in P(E)$, $q_i \in S(E)$ $(1 \leqq 1 \leqq r)$ such that $q_1, \cdots, q_r$ are linearly independent over $\boldsymbol{C}$ and

$$\xi \equiv \sum_i \alpha(p_i) \otimes \beta(\partial(q_i)) \bmod \mathfrak{U}(E).$$

Then

$$\phi_B(\xi) = \sum_i p_i \partial(q_i)$$

and therefore if ξ lies in the kernel of ϕ_B, we can conclude, from Corollary 4 of Lemma 4, that $p_1 = \cdots = p_r = 0$. Hence $\xi \in \mathfrak{U}(E)$ and this proves our assertion.

Now suppose E is the direct sum of two subspaces E_1 and E_2. Let π_i denote the projection of E on E_i $(i = 1, 2)$ corresponding to this sum. Then π_i can be extended uniquely to a homomorphism of $S(E)$ onto $S(E_i)$. We call π_i the projection of $S(E)$ on $S(E_i)$. Note that $S(E_i) \subset S(E)$ (see [2(e), § 6] and $\pi_i p = p$ for $p \in S(E_i)$.

Consider an element $\xi \in \mathfrak{D}(E_1)$. Then

$$\xi = \sum_{1 \leqq i \leqq r} p_i \partial(q_i)$$

where $p_i \in P(E_1)$, $q_i \in S(E_1)$. Since $P(E_1) \subset P(E)$, $S(E_1) \subset S(E)$ [2(e), § 6], the right side can also be interpreted as an element of $\mathfrak{D}(E)$. We claim that in this way, we get an isomorphism of $\mathfrak{D}(E_1)$ into $\mathfrak{D}(E)$. Let $n_1 = \dim E_1$ and assume that the above base $(X_1, \cdots, X_n)$ of E is so chosen that $(X_1, \cdots, X_{n_1})$ is a base for E_1 while the remaining elements form a base for E_2. Define $\mathfrak{X}(E_1)$ and $\mathfrak{U}(E_1)$ corresponding to E_1 as above. Then since $E_1' + \partial(E_1)$ is a subspace of $E' + \partial(E)$, we can identify $\mathfrak{X}(E_1)$ with the subalgebra of $\mathfrak{X}(E)$ generated by 1 and $E_1' + \partial(E_1)$. In order to prove our contention, it would be enough to verify that $\mathfrak{U}(E_1) = \mathfrak{X}(E_1) \cap \mathfrak{U}(E)$. It is obvious that $\mathfrak{U}(E) \supset \mathfrak{U}(E_1)$. So now fix an element $\xi \in \mathfrak{X}(E_1) \cap \mathfrak{U}(E)$. Then we can choose $p_i \in P(E_1)$, $q_i \in S(E_1)$ $(1 \leqq i \leqq r)$ such that $q_1, \cdots, q_r$ are linearly independent over $\boldsymbol{C}$ and

$$\xi \equiv \sum_i \alpha(p_i) \otimes \beta(\partial(q_i)) \bmod \mathfrak{U}(E_1).$$

Then

$$0 = \phi_E(\xi) = \sum_i p_i \partial(q_i)$$

and therefore $p_1 = \cdots = p_r = 0$ as before. This proves that $\xi \in \mathfrak{U}(E_1)$.

Thus $\mathfrak{D}(E_i)$ can be identified with the subalgebra of $\mathfrak{D}(E)$ generated by $(1, E_i' + \partial(E_i))$ under the homomorphism which preserves every element of $E_i' + \partial(E_i)$ $(i = 1, 2)$.

Since $\partial(X_i)\lambda_j = \delta_{ij}$, it is obvious that any element of $\mathfrak{D}(E_1)$ commutes with every element of $\mathfrak{D}(E_2)$. Hence we can define a homomorphism ψ of $\mathfrak{D}(E_1) \otimes \mathfrak{D}(E_2)$ into $\mathfrak{D}(E)$ such that $\psi(D_1 \otimes D_2) = D_1 D_2$ $(D_i \in \mathfrak{D}(E_i)$, $i = 1, 2)$. We shall now prove that ψ is bijective. Since

$$E' + \partial(E) = \sum_{i=1,2} (E_i' + \partial(E_i)),$$

it is obvious that ψ is surjective. Let $\mathfrak{B}$ be the ideal in $\mathfrak{X}(E)$ generated by elements of the form $a \otimes b - b \otimes a$ $(a \in \mathfrak{X}(E_1), b \in \mathfrak{X}(E_2))$. Since $\phi_E(\mathfrak{X}(E_i)) = \mathfrak{D}(E_i)$ $(i = 1, 2)$, it follows from Lemma 5 that $\mathfrak{B} \subset \mathfrak{U}(E)$. Consider the subspace $\mathfrak{X}_{12} = \mathfrak{X}(E_1) \otimes \mathfrak{X}(E_2)$ of $\mathfrak{X}(E)$. Then it is obvious that

$$\mathfrak{X}(E) = \mathfrak{X}_{12} + \mathfrak{B}.$$

We claim that

$$\mathfrak{U}(E) \cap \mathfrak{X}_{12} = \mathfrak{U}_{12}$$

where $\mathfrak{U}_{12} = \mathfrak{X}(E_1) \otimes \mathfrak{U}(E_2) + \mathfrak{U}(E_1) \otimes \mathfrak{X}(E_2)$. It is clear that $\mathfrak{U}_{12}$ is contained in the left side. Now fix $\xi \in \mathfrak{X}_{12}$. Then, in view of what we have seen above, we can choose $p_{ij} \in P(E_1)$, $q_j \in S(E_1)$, $p_{ik}' \in P(E_2)$, $q_k' \in S(E_2)$ $(1 \leqq i \leqq r, 1 \leqq j \leqq s, 1 \leqq k \leqq t)$ such that $(q_1, \cdots, q_s)$ and $(q_1', \cdots, q_t')$ are separately linearly independent over $\boldsymbol{C}$ and

$$\xi \equiv \sum_{i,j,k} \alpha(p_{ij}) \otimes \beta(\partial(q_j)) \otimes \alpha(p_{ik}') \otimes \beta(\partial(q_k')) \bmod \mathfrak{U}_{12}.$$

Therefore

$$\phi_B(\xi) = \sum_{i,j,k} p_{ij}\partial(q_j) \circ p_{ik}'\partial(q_k') = \sum_{i,j,k} p_{ij}p_{ik}'\partial(q_jq_k')$$

since p_{ik}' and $\partial(q_j)$ commute. But the homomorphism of $S(E_1) \otimes S(E_2)$ into $S(E)$ which maps $q \otimes q'$ onto qq' $(q \in S(E_1),\ q' \in S(E_2))$ is obviously bijective. Hence q_jq_k' $(1 \leqq j \leqq s, 1 \leqq k \leqq t)$ are linearly independent in $S(E)$. Therefore if $\phi_B(\xi) = 0$, it follows that

$$\sum_i p_{ij}p_{ik}' = 0$$

for every j, k. But this implies that

$$\sum_i p_{ij} \otimes p_{ik}' = 0$$

in $P(E_1) \otimes P(E_2)$ (see [2(e), § 6]. Therefore

$$\sum_i \alpha(p_{ij}) \otimes b \otimes \alpha(p_{ik}') \otimes c = 0$$

for any $b, c \in \mathfrak{X}(E)$. Therefore $\xi \equiv 0 \bmod \mathfrak{U}_{12}$ and this proves our assertion.

Now suppose $D \in \mathfrak{D}(E_1) \otimes \mathfrak{D}(E_2)$ and $\psi(D) = 0$. We have to show that $D = 0$. Choose $\xi_i \in \mathfrak{D}(E_1), \eta_i \in \mathfrak{D}(E_2)$ $(1 \leqq i \leqq r)$ such that $D = \sum_i \xi_i \otimes \eta_i$. Select $a_i \in \mathfrak{X}(E_1)$, $b_i \in \mathfrak{X}(E_2)$ such that $\phi_B(a_i) = \xi_i$, $\phi_B(b_i) = \eta_i$ $(1 \leqq i \leqq r)$. Then

$$\nu = \sum_i a_i \otimes b_i \in \mathfrak{X}_{12}$$

and

$$\phi_B(\nu) = \sum_i \xi_i\eta_i = \psi(D) = 0.$$

Therefore $\nu \in \mathfrak{U}_{12}$. Put $\phi_{12} = \phi_{B_1} \otimes \phi_{B_2}$. Then ϕ_{12} is a homomorphism of $\mathfrak{X}_{12}$ onto $\mathfrak{D}(E_1) \otimes \mathfrak{D}(E_2)$ and it is obvious that $\mathfrak{U}_{12}$ is contained in the kernel of ϕ_{12}. Hence

$$D = \sum_i \xi_i \otimes \eta_i = \phi_{12}(\nu) = 0.$$

Let $U \supset V$ be two open, nonempty and connected subsets of E. It is easy to see that there exists an injective homomorphism $D \to D_V$ of $\mathfrak{D}_m(U)$ into $\mathfrak{D}_m(V)$ such that g_V is the restriction of g on V $(g \in \mathfrak{M}(U))$ and $(\partial(X))_V$ is just the derivation $\partial(X)$ of $\mathfrak{M}(V)$ $(X \in E)$. We may therefore identify $\mathfrak{D}_m(U)$ with a subalgebra of $\mathfrak{D}_m(V)$ under this homomorphism. Then it is obvious that $\mathfrak{D}_k(U) \subset \mathfrak{D}_k(V)$. Hence, in particular, $\mathfrak{D}(E) \subset \mathfrak{D}_k(E) \subset \mathfrak{D}_k(U)$.

Fix an element $D \in \mathfrak{D}_h(U)$ and a point $X_0 \in U$. Then $D = \sum_{1 \leqq i \leqq r} f_i \partial(q_i)$ where $f_i \in \mathfrak{H}(U)$ and $q_i \in S(E)$. Put $q = \sum_i f_i(X_0) q_i$. Then it is obvious that $f(X_0; D) = f(X_0; \partial(q))$ for every $f \in \mathfrak{H}(U)$. Taking into account Corollary 2 of Lemma 4, we obtain the following result. There exists a unique element $D_{X_0} \in \partial(S(E))$ such that $f(X_0; D) = f(X_0; D_{X_0})$ for all $f \in \mathfrak{H}(U)$. D_{X_0} is called the local expressions of D at X_0.

Now fix $D \in \mathfrak{D}_m(U)$ and $X_0 \in U$. We say that D is *defined* at X_0, if we can choose an open connected neighborhood V of X_0 in U such that $D \in \mathfrak{D}_h(V)$. If this is so, we again denote by D_{X_0} the local expression of D at X_0. (The value of D_{X_0} is independent of the choice of V.)

Let $Q(E)$ denote the quotient field of $P(E)$. Then $Q(E) \subset \mathfrak{M}(E)$. Let $\mathfrak{D}_r(E)$ be the subalgebra of $\mathfrak{D}_m(E)$ generated by $Q(E)$ and $\mathfrak{D}(E)$. We call $\mathfrak{D}_r(E)$ the algebra of rational differential operators on E.

LEMMA 6. *Let D be an element in $\mathfrak{D}_r(E)$ and X_0 a point in E. Suppose D is defined at X_0. Then we can choose $p \in P(E)$ such that $p(X_0) \neq 0$ and $pD \in \mathfrak{D}(E)$. Moreover*

$$D_{X_0} = p(X_0)^{-1}(pD)_{X_0}$$

for any such p.

Choose an open connected neighborhood U of X_0 in E such that $D \in \mathfrak{D}_h(U)$. Since $\mathfrak{D}_r(E) = Q(E)\mathfrak{D}(E)$, we can choose $p_0 \in P(E)$ $(p_0 \neq 0)$ such that $p_0 D \in \mathfrak{D}(E)$. Therefore we can select $q_i \in S(E)$, $f_i \in \mathfrak{H}(U)$ and $p_i \in P(E)$ $(1 \leqq i \leqq r)$ such that $q_1, \cdots, q_r$ are linearly independent over $\boldsymbol{C}$ and

$$p_0 D = \sum_{1 \leqq i \leqq r} p_i \partial(q_i),$$

$$D = \sum_{1 \leqq i \leqq r} f_i \partial(q_i),$$

where the second equation is meant to hold in $\mathfrak{D}_h(U)$. Then it is clear that

$$(p_0 D)_X = \sum_i p_0(X) f_i(X) \partial(q_i) = \sum_i p_i(X) \partial(q_i) \qquad (X \in U).$$

Therefore $p_i = p_0 f_i$ $(1 \leqq i \leqq r)$ on U. This implies that X_0 does not lie on the polar divisor of any one of the rational functions p_i/p_0 $(1 \leqq i \leqq r)$. Let q denote the h. c. f. of $(p_0, p_1, \cdots, p_r)$ in $P(E)$ aid put $p = p_0/q$. Then $p \in P(E)$ and it is clear that $pD \in \mathfrak{D}(E)$ and X_0 does not lie in the divisor of zeros of p. Therefore $p(X_0) \neq 0$. The second statement of the lemma is obvious.

The group $\boldsymbol{GL}(E)$ operates on E. If $f \in \mathfrak{H}(E)$ and $x \in \boldsymbol{GL}(E)$ we denote by f^x the function $X \rightarrow f(x^{-1}X)$ $(X \in E)$. Then $f \rightarrow f^x$ is an automorphism of $\mathfrak{H}(E)$ which can be extended uniquely to an automorphism of $\mathfrak{M}(E)$ over $\boldsymbol{C}$. For any $D \in \mathfrak{D}_m(E)$, define D^x by

$$D^x g = (D g^{x^{-1}})^x \qquad (g \in \mathfrak{M}(E)).$$

Then $D \rightarrow D^x$ is an automorphism of $\mathfrak{D}_m(E)$ which leaves $P(E)$ and $Q(E)$ invariant. Also $(\partial(X))^x = \partial(xX)$ $(X \in E)$. Hence x leaves $\partial(S(E))$, $\mathfrak{D}(E)$ and $\mathfrak{D}_r(E)$ invariant.

LEMMA 7. *Let $D \in \mathfrak{D}_r(E)$, $x \in \boldsymbol{GL}(E)$ and $X \in E$. Suppose D is defined at X. Then D^x is defined at xX and $(D^x)_{xX} = (D_X)^x$.*

Choose $p \in P(E)$ such that $p(X) \neq 0$ and $\xi = pD \in \mathfrak{D}(E)$. Then $D = p^{-1}\xi$ and therefore $D^x = (p^x)^{-1}\xi^x$. Now $p^x D^x = \xi^x \in \mathfrak{D}(E)$ and $p^x(xX) = p(X) \neq 0$. Hence D^x is defined at xX. Moreover

$$(D^x)_{xX} = (p^x(xX))^{-1}(\xi^x)_{xX} = p(X)^{-1}(\xi^x)_{xX}.$$

Now fix $q \in P(E)$. Then $\xi^x q = (\xi q^{x^{-1}})^x$ and $(\xi_X)^x q = (\xi_X q^{x^{-1}})^x$. Hence

$$q(Y;\xi^x) = q^{x^{-1}}(x^{-1}Y;\xi), \qquad q(Y;(\xi_X)^x) = q^{x^{-1}}(x^{-1}Y;\xi_X)$$

for $Y \in E$. Putting $Y = xX$, we get

$$q(xX;(\xi^x)_{xX}) = q^{x^{-1}}(X;\xi) = q(xX;(\xi_X)^x).$$

This being true for every $q \in P(E)$, we conclude from Corollary 2 of Lemma 4 that $(\xi^x)_{xX} = (\xi_X)^x$. Hence

$$(D^x)_{xX} = (p(X)^{-1}\xi_X)^x = (D_X)^x.$$

COROLLARY. *Let U be a subset of E with the following two properties.*

1) *If $p \in P(E)$ and p vanishes identically on U then $p = 0$.*

2) *D is defined at every point of $U \cup x^{-1}U$.*

Then if $D_{xX} = (D_X)^x$ for $X \in U$, we can conclude that $D = D^x$.

It follows from the above lemma that both D and D^x are defined at every point $Y \in U$ and $(D - D^x)_Y = 0$. Choose $p \in P(E)$ $(p \neq 0)$ such that $\Delta = p(D - D^x) \in \mathfrak{D}(E)$. Then if $q \in P(E)$,

$$q(Y;\Delta) = q(Y;\Delta_Y) = 0$$

for $Y \in U$. Hence it follows from condition 1) that $\Delta q = 0$. But then we can conclude from Corollary 3 of Lemma 4 that $\Delta = 0$ and therefore $D - D^x = 0$.

LEMMA 8. *Fix an element* $\zeta \neq 0$ *in* $Q(E)$. *Then there exists a unique automorphism* μ_ζ *of* $\mathfrak{D}_r(E)$ *such that* $\mu_\zeta(q) = q$ *for* $q \in Q(E)$ *and*

$$\mu_\zeta(\partial(X)) = \partial(X) + (1/2)\zeta^{-1}(\partial(X)\zeta) \qquad (X \in E).$$

Moreover $(\mu_\zeta(D))^x = \mu_{\zeta^x}(D^x)$ $(x \in \boldsymbol{GL}(E), D \in \mathfrak{D}_r(E))$ *and*

$$\mu_{\zeta_1\zeta_2} = \mu_{\zeta_1}\mu_{\zeta_2} \qquad (\zeta_i \in Q(E), \zeta_i \neq 0, i = 1, 2).$$

Let U be a non-empty open convex set in E such that U does not meet the divisor of ζ. Fix a point $X_0 \in U$ and define

$$\log \zeta(X) = c + \int_{X_0}^{X} \zeta^{-1}\, d\zeta \qquad (X \in U)$$

where the integral is taken along the straight line-segment from X_0 to X and c is a constant such that $\zeta(X_0) = e^c$. Put

$$\eta = \exp((1/2)\log\zeta).$$

Then η is a holomorphic function on U and $\zeta = \eta^2$ on U. Let ν_η denote the automorphism of $\mathfrak{D}_m(U)$ given by

$$\nu_\eta(D) = \eta^{-1}D \circ \eta \qquad (D \in \mathfrak{D}_m(U)).$$

It is clear that

$$\begin{aligned}\nu_\eta(\partial(X)) &= \eta^{-1}\partial(X) \circ \eta \\ &= \partial(X) + \eta^{-1}(\partial(X)\eta) \\ &= \partial(X) + (1/2)\zeta^{-1}(\partial(X)\zeta)\end{aligned}$$

for $X \in E$. Moreover $\nu_\eta(q) = q$ for $q \in Q(E)$. Hence ν_η obviously defines an automorphism of $\mathfrak{D}_r(E)$. Since the uniqueness of μ_ζ is clear, this proves the first statement of the lemma.

Let α_1, α_2 be two automorphisms of $\mathfrak{D}_r(E)$. If $\alpha_1(\partial(X)) = \alpha_2(\partial(X))$ $(X \in E)$ and $\alpha_1(q) = \alpha_2(q)$ $(q \in Q(E))$, then it is obvious that $\alpha_1 = \alpha_2$. The remaining statements of the lemma follow immediately from this fact.

Remark. We shall usually write $\zeta^{-\frac{1}{2}}D \circ \zeta^{\frac{1}{2}}$ instead of $\mu_\zeta(D)$ $(D \in \mathfrak{D}_r(E))$. It is clear from the above proof that if D, ζ and ζ^{-1} are all defined at a point $X_0 \in E$, then $\zeta^{\frac{1}{2}}D \circ \zeta^{\frac{1}{2}}$ is also defined at X_0.

4. The main theorem on invariant differential operators. Let $\mathfrak{g}$ be a reductive Lie algebra over $\boldsymbol{C}$ and G its connected adjoint group. Then $G \subset \boldsymbol{GL}(\mathfrak{g})$. Let $\mathfrak{J}_r(\mathfrak{g})$ be the subalgebra of those elements $D \in \mathfrak{D}_r(\mathfrak{g})$ for which $D^x = D$ $(x \in G)$.

LEMMA 9. *For any $X \in \mathfrak{g}$, let $\tau(X)$ denote the element of $\mathfrak{D}(\mathfrak{g})$ given by*

$$(\tau(X))_Y = \partial([Y, X]) \qquad (Y \in \mathfrak{g}).$$

Then if $D \in \mathfrak{D}_r(\mathfrak{g})$, the following two conditions on D are equivalent.

1) $D \in \mathfrak{I}_r(\mathfrak{g})$.

2) $\tau(X) \circ D - D \circ \tau(X) = 0$ *for all* $X \in \mathfrak{g}$.

Fix $p \in P(\mathfrak{g})$ $(p \neq 0)$ such that $\xi = pD$ lies in $\mathfrak{D}(\mathfrak{g})$. Let U_0 be a non-empty open set in $\mathfrak{g}$ such that p is nowhere zero on U_0. Fix $Y_0 \in U_0$ and choose open connected neighborhoods U and V of Y_0 and 1 in U_0 and G respectively such that $V^{-1} = V$ and $xY \in U_0$ for $x \in V$ and $Y \in U$. Then if $q \in P(\mathfrak{g})$, the functions $(x, Y) \rightarrow (Dq^x)(Y)$ and $(x, Y) \rightarrow (Dq)^x(Y)$ are both holomorphic on $V \times U$.

Now suppose $D \in \mathfrak{I}_r(\mathfrak{g})$. Fix $q \in P(\mathfrak{g})$ and $X \in \mathfrak{g}$. Then

$$(Dq)^{\exp tX}(Y) = (Dq^{\exp tX})(Y) \qquad (Y \in U)$$

if t is sufficiently near zero in $\boldsymbol{C}$. From this it follows that $D_1 q = 0$ on U. Here $D_1 = \tau(X) \circ D - D \circ \tau(X)$. But $D_1 \in \mathfrak{D}_r(\mathfrak{g})$ and therefore $D_1 q \in Q(\mathfrak{g})$. Hence $D_1 q = 0$. But then we can conclude from Corollary 3 of Lemma 4 that $D_1 = 0$.

Conversely suppose 2) holds. Then the proof of Lemma 15 of [2(e)] is applicable and we conclude that

$$D_{xX} = (D_X)^x \qquad (x \in V, X \in U).$$

But then $D = D^x$ $(x \in V)$ from the corollary of Lemma 7. Since G is connected, V generates G and therefore it follows that $D \in \mathfrak{I}_r(\mathfrak{g})$.

Now we shall use the notation of § 2 (with $k = \boldsymbol{C}$). In particular $\mathfrak{z}$ and ζ have the same meaning as before. Let D be an element of $\mathfrak{D}_r(\mathfrak{g})$ which is defined at some point of $\mathfrak{z}$. Then we shall associate to it an element $\Delta \in \mathfrak{D}_r(\mathfrak{g})$ as follows. Choose $q \in P(\mathfrak{g})$ such that 1) q does not vanish identically on $\mathfrak{z}$ and 2) $D' = qD$ is in $\mathfrak{D}(\mathfrak{g})$. This is possible from Lemma 6. For any $Z \in \mathfrak{z}$, define $p_Z \in S(\mathfrak{g})$ by the condition $(D')_Z = \partial(p_Z)$. Then it follows from Corollary 2 of Lemma 2, that there exists an element $\Delta' \in \mathfrak{D}_r(\mathfrak{z})$ which is defined everywhere on $\mathfrak{z}'$ and

$$(\Delta')_Z = \partial(\alpha_Z(p_Z)) \qquad (Z \in \mathfrak{z}')$$

We define $\Delta = q_{\mathfrak{z}}^{-1}\Delta'$ where $q_{\mathfrak{z}}$ denotes the restriction of q on $\mathfrak{z}$. It is easy to verify that Δ does not depend on the choice of q. It is also clear from Lemma

6 that Δ is defined at all points of $\mathfrak{z}'$ at which D is defined. We shall denote Δ by $\delta_{\mathfrak{g}/\mathfrak{z}}'(D)$.

Let $\mathfrak{I}(\mathfrak{g}) = \mathfrak{D}(\mathfrak{g}) \cap \mathfrak{I}_r(\mathfrak{g})$.

LEMMA 10. *If $D \in \mathfrak{I}(\mathfrak{g})$, then $\delta_{\mathfrak{g}/\mathfrak{z}}'(D)$ is invariant under the normalizer of $\mathfrak{z}$ in G.*

Let $\Delta = \delta_{\mathfrak{g}/\mathfrak{z}}'(D)$. Then Δ is defined everywhere on $\mathfrak{z}'$. Fix $x \in G$ such that $x\mathfrak{z} = \mathfrak{z}$. Then by Lemma 3, $x\mathfrak{z}' = \mathfrak{z}'$. So, in view of the corollary of Lemma 7, it would be enough to prove that

$$(\Delta_Z)^x = \Delta_{xZ} \qquad (Z \in \mathfrak{z}').$$

For $X \in \mathfrak{g}$, define $p_X \in S(\mathfrak{g})$ by the condition $D_X = \partial(p_X)$. Then $p_{xX} = (p_X)^x$ since $D \in \mathfrak{I}(\mathfrak{g})$ (Lemma 7). Hence

$$\begin{aligned} \Delta_{xZ} &= \partial(\alpha_{xZ}(p_{xZ})) = \partial((\alpha_Z(p_Z))^x) \\ &= (\Delta_Z)^x \qquad (Z \in \mathfrak{z}') \end{aligned}$$

from Lemma 3.

Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{z}$ and $\mathfrak{h}'$ the set of those elements in $\mathfrak{h}$ which are regular in $\mathfrak{g}$. It is obvious that $\mathfrak{h}' \subset \mathfrak{z}'$. Now fix $D \in \mathfrak{D}(\mathfrak{g})$. Then $\Delta = \delta_{\mathfrak{g}/\mathfrak{z}}'(D)$ is defined everywhere on $\mathfrak{z}'$ and it is easy to see that the element $\delta_{\mathfrak{g}/\mathfrak{h}}'(\Delta)$ of $\mathfrak{D}_r(\mathfrak{h})$ is defined everywhere on $\mathfrak{h}'$.

LEMMA 11. *$\delta_{\mathfrak{g}/\mathfrak{h}}'(D) = \delta_{\mathfrak{z}/\mathfrak{h}}'(\delta_{\mathfrak{g}/\mathfrak{z}}'(D))$ for $D \in \mathfrak{D}(\mathfrak{g})$.*

Fix a point $H \in \mathfrak{h}'$ and let $\Delta = \delta_{\mathfrak{g}/\mathfrak{z}}'(D)$ and $\Box = \delta_{\mathfrak{z}/\mathfrak{h}}'(\Delta)$. Choose $p \in S(\mathfrak{g})$ such that $D_H = \partial(p)$. Then $\Delta_H = \partial(\alpha_H(p))$. Now

$$p - \alpha_H(p) \in \Gamma_H(\mathfrak{Q}_+ \otimes S(\mathfrak{z}))$$

from the definition of $\alpha_H(p)$. Moreover if $\Box_H = \partial(q)$ $(q \in S(\mathfrak{h}))$, we have

$$\alpha_H(p) - q \in \Gamma_H(\mathfrak{G}_+ \otimes S(\mathfrak{h})).$$

On the other hand if we apply Lemma 2 to $(\mathfrak{z}, \mathfrak{h})$ instead of $(\mathfrak{g}, \mathfrak{h})$, we get

$$S(\mathfrak{z}) = \Gamma_H(\mathfrak{S}(\mathfrak{z}) \otimes S(\mathfrak{h})).$$

Therefore

$$p - \alpha_H(p) \in \Gamma_H(\mathfrak{Q}_+\mathfrak{S}(\mathfrak{z}) \otimes S(\mathfrak{h})) \subset \Gamma_H(\mathfrak{G}_+ \otimes S(\mathfrak{h})).$$

This proves that

$$p - q = (p - \alpha_H(p)) + (\alpha_H(p) - q) \in \Gamma_H(\mathfrak{G}_+ \otimes S(\mathfrak{h})).$$

Hence if we take into account the remark after Lemma 2, it becomes clear that $(\delta_{\mathfrak{g}/\mathfrak{h}}'(D))_H = \partial(q) = \square_H$. This proves that $\delta_{\mathfrak{g}/\mathfrak{h}}'(D) = \square$.

Fix a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ and let W denote the Weyl group of $(\mathfrak{g}, \mathfrak{h})$. Let $\mathfrak{I}_r(\mathfrak{h})$ denote the subalgebra of those elements of $\mathfrak{D}_r(\mathfrak{h})$ which are invariant under W. Put $\mathfrak{I}(\mathfrak{h}) = \mathfrak{D}(\mathfrak{h}) \cap \mathfrak{I}_r(\mathfrak{h})$, $J(\mathfrak{g}) = P(\mathfrak{g}) \cap \mathfrak{I}(\mathfrak{g})$ and $J(\mathfrak{h}) = P(\mathfrak{h}) \cap \mathfrak{I}(\mathfrak{h})$. Also let $\pi = \prod_{\alpha > 0} \alpha$ where α runs over all roots of $(\mathfrak{g}, \mathfrak{h})$ which are positive (under some fixed order). Then $\pi \in P(\mathfrak{h})$.

THEOREM 1.[2] *Put*

$$\delta_{\mathfrak{g}/\mathfrak{h}}(D) = \pi \delta_{\mathfrak{g}/\mathfrak{h}}'(D) \circ \pi^{-1}$$

for $D \in \mathfrak{I}(\mathfrak{g})$. *Then* $\delta_{\mathfrak{g}/\mathfrak{h}}(D) \in \mathfrak{I}(\mathfrak{h})$ *and* $\delta_{\mathfrak{g}/\mathfrak{h}}$ *is a homomorphism of* $\mathfrak{I}(\mathfrak{g})$ *into* $\mathfrak{I}(\mathfrak{h})$. *Moreover an element* D *of* $\mathfrak{I}(\mathfrak{g})$ *lies in the kernel of this homomorphism if and only if* $Dp = 0$ *for all* $p \in J(\mathfrak{g})$.

We shall write δ and δ' for $\delta_{\mathfrak{g}/\mathfrak{h}}$ and $\delta_{\mathfrak{g}/\mathfrak{h}}'$ respectively. It is obvious that $\delta(D)$ lies in $\mathfrak{D}_r(\mathfrak{h})$ and its definition is independent of the order of roots.

Let $\mathfrak{c}$ be the center and $\mathfrak{g}_1$ the derived algebra of $\mathfrak{g}$ so that $\mathfrak{g} = \mathfrak{c} + \mathfrak{g}_1$, where the sum is direct. Then $\mathfrak{c} \subset \mathfrak{h}$ and $\mathfrak{h} = \mathfrak{c} + \mathfrak{g}_1 \cap \mathfrak{h}$. We can choose a compact real form $\mathfrak{g}_{10}$ of $\mathfrak{g}_1$ such that $\mathfrak{g}_{10} \cap \mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}_1$ (see, for example [3, p. 155]). Choose a real subspace $\mathfrak{c}_0$ of $\mathfrak{c}$ such that $\dim_R \mathfrak{c}_0 = \dim_C \mathfrak{c}$ and $\mathfrak{c}_0$ spans $\mathfrak{c}$ over $\boldsymbol{C}$. Then $\mathfrak{g}_0 = \mathfrak{c}_0 + \mathfrak{g}_{10}$ is a real form of $\mathfrak{g}$ and $\mathfrak{h}_0 = \mathfrak{h} \cap \mathfrak{g}_0$ is a Cartan subalgebra of $\mathfrak{g}_0$. Let $\mathfrak{h}'$ be the set of all points H in $\mathfrak{h}$ where $\pi(H) \neq 0$. Then $\delta(D)$ is defined everywhere on $\mathfrak{h}'$. It is clear that if $q \in P(\mathfrak{h})$ and $q = 0$ on $\mathfrak{h}_0' = \mathfrak{h}_0 \cap \mathfrak{h}'$, then $q = 0$. Hence in order to prove that δ is a homomorphism of $\mathfrak{I}(\mathfrak{g})$ into $\mathfrak{D}_r(\mathfrak{h})$, it is enough to verify that

$$(\delta(D_1 D_2) - \delta(D_1) \circ \delta(D_2))_H = 0$$

for $H \in \mathfrak{h}_0'$ and $D_1, D_2 \in \mathfrak{I}(\mathfrak{g})$. This is proved in the same way as Lemma 7 of [2(a)].

For a given $s \in W$, we can choose x in G such that $\mathrm{Ad}(x) = s$ on $\mathfrak{h}$. Then if $D \in \mathfrak{I}(\mathfrak{g})$, it follows from Lemma 10 that

$$(\delta'(D))^s = (\delta'(D))^x = \delta'(D).$$

Moreover $\pi^s = \epsilon(s)\pi$ where $\epsilon(s) = 1$ or -1. Therefore it is clear that $\delta(D) \in \mathfrak{I}_r(\mathfrak{h})$. Hence in order to prove that $\delta(D) \in \mathfrak{I}(\mathfrak{h})$, it would be enough to verify that $\delta(D) \in \mathfrak{D}(\mathfrak{h})$.

---

[2] This theorem is a generalization of Theorem 1 of [2(a)].

We know (see § 3) that $\mathfrak{D}(\mathfrak{g}) = \mathfrak{D}(\mathfrak{c})\mathfrak{D}(\mathfrak{g}_1)$. Hence $D = \sum_{1 \leqq i \leqq r} \xi_i \eta_i$ where $\xi_i \in \mathfrak{D}(\mathfrak{c})$, $\eta_i \in \mathfrak{D}(\mathfrak{g}_1)$ and $\xi_1, \cdots, \xi_r$ are linearly independent over $\boldsymbol{C}$. Then since $D \in \mathfrak{J}(\mathfrak{g})$, it is obvious that $\eta_i \in \mathfrak{J}(\mathfrak{g}_1)$. Moreover $\delta(\xi) = \xi$ for $\xi \in \mathfrak{D}(\mathfrak{c})$. Hence it follows from what we have proved above, that

$$\delta(D) = \sum_i \xi_i \delta(\eta_i).$$

So it would be sufficient to prove that $\delta(\eta_i) \in \mathfrak{D}(\mathfrak{h}_1)$ where $\mathfrak{h}_1 = \mathfrak{h} \cap \mathfrak{g}_1$. Thus in order to complete the proof of the first part of Theorem 1 we can confine ourselves to the case when $\mathfrak{g}$ is semisimple. This will be done in § 6.

We now come to the second assertion of the theorem. Fix $p \in J(\mathfrak{g})$ and $D \in \mathfrak{J}(\mathfrak{g})$. Then it follows from the defintion of $\delta'(D)$ that $p(H;D) = p(H;\delta'(D))$ for $H \in \mathfrak{h}'$ (see [2(a), Lemma 6]). Therefore if $\delta(D) = 0$, it is obvious that $Dp = 0$ on $\mathfrak{h}'$. Let $n = \dim \mathfrak{g}$. Then it is easy to verify that the holomorphic mapping $(x, H) \rightarrow xH$ is of rank n on $G \times \mathfrak{h}'$. Therefore $(\mathfrak{h}')^G$ is open in $\mathfrak{g}$ and hence $Dp = 0$. Conversely suppose $\delta(D) \neq 0$. It follows from Corollary 2 of Lemma 2 that we can choose an integer $m \geqq 0$ such that $\nabla = \pi^{2m}\delta'(D) \in \mathfrak{J}(\mathfrak{h})$. It is obvious that $\nabla \neq 0$ and so we can select a point $H_0 \in \mathfrak{h}_0'$ such that $\nabla_H \neq 0$. Since $\mathfrak{g}_{10}$ is a *compact* real form of $\mathfrak{g}_1$, $\mathfrak{h}_0$ is stable under W. Moreover since H_0 is regular in $\mathfrak{g}$, $sH_0 \neq H_0$ for $s \neq 1$ in W. Fix an open neighborhood U of H_0 in $\mathfrak{h}_0'$ such that $sU \cap U = \emptyset$ for $s \neq 1$ in W. Then we can choose $f_1 \in C_c^\infty(U)$ such that $f_1(H_0;\nabla) = f_1(H_0;\nabla_{H_0}) \neq 0$. Put $f(H) = \sum_{s \in W} f_1(sH)$ $(H \in \mathfrak{h}_0)$. Then f is invariant under W and $f(H_0;\nabla) = f_1(H_0;\nabla) \neq 0$. On the other hand we can select $q_1 \in P(\mathfrak{h})$ such that

$$|f(sH_0;\nabla) - q_1(sH_0;\nabla)| \leqq (1/2)|f(H_0;\nabla)|$$

for $s \in W$. Put

$$q = [W]^{-1} \sum_{s \in W} q_1^s$$

where $[W]$ denotes the order of W. Since f and ∇ are both invariant under W, it is clear that

$$f(H_0;\nabla) - q(H_0;\nabla) = [W]^{-1} \sum_{s \in W} \{f(sH_0;\nabla) - q_1(sH_0;\nabla)\}$$

and therefore

$$|f(H_0;\nabla) - q(H_0;\nabla)| \leqq (1/2)|f(H_0;\nabla)|.$$

This shows that $q(H_0;\nabla) \neq 0$. But $q \in J(\mathfrak{h})$ and therefore by Chevalley's Theorem (see [2(a), Lemma 9]), there exists an element $p \in J(\mathfrak{g})$ whose restriction on $\mathfrak{h}$ is q. Therefore

$$p(H_0;D) = q(H_0;\delta'(D)) = \pi(H_0)^{-2m} q(H_0;\nabla) \neq 0$$

and this proves that $Dp \neq 0$.

5. A simple lemma. Let E be a vector space over $\boldsymbol{C}$ of finite dimension. Assume that we are given a nondegenerate symmetric bilinear form B on E. We identify E with its dual under B. Let $(-1)^{\frac{1}{2}}$ denote a fixed square-root of -1.

LEMMA 12. *There exists a unique automorphism $D \rightarrow \hat{D}$ of $\mathfrak{D}(E)$ such that*

$$\hat{X} = -(-1)^{\frac{1}{2}}\partial(X),\ (\partial(X))^{\wedge} = -(-1)^{\frac{1}{2}}X \qquad (X \in E).$$

Similarly there exists a unique anti-automorphism $D \rightarrow D^*$ of $\mathfrak{D}(E)$ such that

$$X^* = X, \qquad \partial(X)^* = -\partial(X) \qquad (X \in E).$$

Let x be an element of $\boldsymbol{GL}(E)$ which leaves B invariant. Then

$$(D^x)^* = (D^*)^x, \qquad (D^x)^{\wedge} = (\hat{D})^x$$

for $D \in \mathfrak{D}(E)$.

We use the notation of § 3. Since $E = E'$, $(E + \partial(E), 1)$ generates $\mathfrak{D}(E)$ and therefore the uniqueness of both $*$ and $\hat{}$ is obvious. We shall now verify their existence. $\mathfrak{T}(E)$ is now the tensor algebra over $E + \partial(E)$. Define two linear mappings $*$ and $\hat{}$ of $E + \partial(E)$ onto itself according to the statement of the lemma. Then obviously $\hat{}$ and $*$ can be extended uniquely to an automorphism and an anti-automorphism respectively of $\mathfrak{T}(E)$. It is easy to verify that the set of generators of $\mathfrak{U}(E)$ (described in § 3) is mapped onto itself by each of these operations. Therefore we get the corresponding operations in $\mathfrak{D}(E) = \mathfrak{T}(E)/\mathfrak{U}(E)$. The last assertion of the lemma follows from the obvious fact that if two automorphisms (or anti-automorphisms) of $\mathfrak{D}(E)$ coincide on $E + \partial(E)$, then they must be identical.

6. Completion of the proof of Theorem 1. Let $\mathfrak{g}$ be a semisimple Lie algebra over $\boldsymbol{C}$ and $\mathfrak{h}$ a Cartan subalgebra of $\mathfrak{g}$. Put $B(X,Y) = \operatorname{tr}(\operatorname{ad} X \operatorname{ad} Y)$ $(X, Y \in \mathfrak{g})$ and let $B_{\mathfrak{h}}$ denote the restrictions of B on $\mathfrak{h}$. Then both B and $B_{\mathfrak{h}}$ are symmetric and non-degenerate and so we can identify $\mathfrak{g}$ and $\mathfrak{h}$ with their dual spaces respectively under these bilinear forms. Put $\delta = \delta_{\mathfrak{g}/\mathfrak{h}}$ (in the notation of Theorem 1) and define $\tau(X)$ $(X \in \mathfrak{g})$ as in Lemma 9.

LEMMA 13. *Let D be an element of $\mathfrak{I}(\mathfrak{g})$. Then $\hat{D}$ and D^* are also in $\mathfrak{I}(\mathfrak{g})$. Moreover $\delta(D) \in \mathfrak{I}(\mathfrak{h})$ and*

$$\delta(\hat{D}) = (\delta(D))\hat{}, \qquad \delta(D^*) = \delta(D)^*.$$

Finally $(\tau(Y))\hat{} = \tau(Y)$ *and* $(\tau(Y))^* = -\tau(Y)$ *for* $Y \in \mathfrak{g}$.

We can choose, as before, a compact real form $\mathfrak{g}_0$ of $\mathfrak{g}$ such that $\mathfrak{h}_0 = \mathfrak{h} \cap \mathfrak{g}_0$ is a Cartan subalgebra of $\mathfrak{g}_0$. Let G_0 be the real analytic subgroup of G corresponding to $\mathfrak{g}_0$. Then G_0 is compact. Define the space $\mathcal{B}(\mathfrak{g}_0)$ as in [2(a), §§2 and 5], and for any $f \in \mathcal{B}(\mathfrak{g}_0)$, put

$$\phi_f(H) = \pi(H) \int_{G_0} f(xH)\, dx \qquad (H \in \mathfrak{h}_0)$$

where the Haar measure dx on G_0 is so normalized as to make the total measure of G_0 equal to 1. Then we know [2(a), Theorem 3] that $\phi_f \in \mathcal{B}(\mathfrak{h}_0)$. As before, let $\mathfrak{h}_0'$ be the set of all regular elements in $\mathfrak{h}_0$. Then if $D \in \mathfrak{D}(\mathfrak{g})$, we know from §4 that $\delta(D) = \pi\delta_{\mathfrak{g}/\mathfrak{h}}'(D) \circ \pi^{-1}$ is defined everywhere on $\mathfrak{h}_0'$ and it follows from the proof of Lemma 15 of [2(a)] that

$$\phi_{Df} = \delta(D)\phi_f$$

on $\mathfrak{h}_0'$. Let dX and dH denote Euclidean measures on $\mathfrak{g}_0$ and $\mathfrak{h}_0$ respectively. We normalize them in such a way that

$$\int_{\mathfrak{h}_0} f(X)\, dX = \int_{\mathfrak{g}_0} |\pi(H)|^2\, dH \int_{G_0} f(xH)\, dx$$

for $f \in C_c(\mathfrak{g}_0)$ (see [2(a), p. 105]). Now for $f \in \mathcal{B}(\mathfrak{g}_0)$ and $g \in \mathcal{B}(\mathfrak{h}_0)$, define their Fourier transforms $\hat{f}$ and $\hat{g}$ by

$$\hat{f}(Y) = \int f(X) \exp((-1)^{\frac{1}{2}} B(X, Y))\, dX \qquad (Y \in \mathfrak{g}_0)$$

and

$$\hat{g}(H') = \int g(H) \exp((-1)^{\frac{1}{2}} B(H, H'))\, dH \qquad (H' \in \mathfrak{h}_0)$$

Then there exists [2(a), Theorem 3] a number $c \neq 0$ such that

$$\phi_{\hat{f}} = c(\phi_f)\hat{}$$

for all $f \in \mathcal{B}(\mathfrak{g}_0)$.

We know [2(a), Lemma 1] that $(Df)\hat{} = \hat{D}\hat{f}$ for $D \in \mathfrak{D}(\mathfrak{g})$ and $f \in \mathcal{B}(\mathfrak{g}_0)$. Similarly D^* is the adjoint of D and corresponding statements hold for elements of $\mathfrak{D}(\mathfrak{h})$.

Since B is invariant under G, it follows immediately that

$$(f^y)\hat{} = (\hat{f})^y$$

for $y \in G_0$ and $f \in \mathcal{B}(\mathfrak{g}_0)$. Moreover it follows from Lemma 12 that the operations of G on $\mathfrak{D}(\mathfrak{g})$ commute with both ^ and *. Hence $\mathfrak{J}(\mathfrak{g})$ is mapped into itself by ^ and *.

Fix $Y \in \mathfrak{g}_0$ and put $y_t = \exp tY$ $(t \in \boldsymbol{R})$. Then

$$f(X;\tau(Y)) = \{df^{y_t}(X)/dt\}_{t=0} = \{df(y_t^{-1}X)/dt\}_{t=0} \qquad (X \in \mathfrak{g}_0)$$

Hence if $f \in C_c^\infty(\mathfrak{g}_0)$, it is obvious that $(\tau(Y)f)\hat{}$ is the function

$$X \rightarrow \{d\hat{f}(y_t^{-1}X)/dt\}_{t=0} = \hat{f}(X;\tau(Y)).$$

This proves that $(\tau(Y))\hat{}\,\hat{f} = \tau(Y)\hat{f}$. Since $C_c^\infty(\mathfrak{g}_0)$ is dense in $\mathscr{B}(\mathfrak{g}_0)$, this shows that $(\tau(Y))\hat{} = \tau(Y)$ for $Y \in \mathfrak{g}_0$ and therefore also for $Y \in \mathfrak{g}$. Moreover we know from Lemma 6 of [2(e)] that $\tau(X)^* = -\tau(X)$ $(X \in \mathfrak{g})$.

Now fix $D \in \mathfrak{J}(\mathfrak{g})$ and first let us assume that $\delta(D) \in \mathfrak{D}(\mathfrak{h})$. Then we have seen in § 4 that $\delta(D) \in \mathfrak{J}(\mathfrak{h})$. But $(Df)\hat{} = \hat{D}\hat{f}$ and therefore

$$\begin{aligned} \phi_{\hat{D}\hat{f}} = \phi_{(Df)\hat{}} = c(\phi_{Df})\hat{} &= c(\delta(D)\phi_f)\hat{} \\ &= c(\delta(D))\hat{}(\phi_f)\hat{} \end{aligned}$$

for $f \in \mathscr{B}(\mathfrak{g}_0)$. Also

$$\phi_{\hat{D}\hat{f}} = \delta(\hat{D})\phi_{\hat{f}} = c\delta(\hat{D})(\phi_f)\hat{}$$

on $\mathfrak{h}_0'$. Let $\mathscr{B}_W(\mathfrak{h}_0)$ denote the set of all functions on $\mathscr{B}(\mathfrak{h}_0)$ which are invariant under W. Then it follows from [2(a), Theorem 4] and the above result that

$$\delta(D)\hat{}(\pi g) = \delta(\hat{D})(\pi g)$$

on $\mathfrak{h}_0'$ for every $g \in \mathscr{B}_W(\mathfrak{h}_0)$. Fix an element $H_0 \in \mathfrak{h}_0'$ and an open neighborhood U of H_0 in $\mathfrak{h}_0'$ such that $sU \cap U = \emptyset$ for $s \neq 1$ in W. For any $\beta \in C_c^\infty(U)$, put

$$g_\beta = \pi^{-1} \sum_{s \in W} \epsilon(s)\beta^s.$$

Then $g_\beta \in C_c^\infty(\mathfrak{h}_0')$ and it is invariant under W. Let ξ be an element of $\mathfrak{J}_r(\mathfrak{h})$ which is defined everywhere on $\mathfrak{h}_0'$. Then it is clear that $\xi\beta^s = (\xi\beta)^s = 0$ on U if $s \neq 1$. Hence if we take $g = g_\beta$, in our result above, we get

$$\delta(D)\hat{}\beta = \delta(\hat{D})\beta$$

on U. This proves that $(\delta(D)\hat{} - \delta(\hat{D}))_{H_0} = 0$. But since $\delta(D)\hat{} - \delta(\hat{D})$ lies in $\mathfrak{D}_r(\mathfrak{h})$ and $\mathfrak{h}_0'$ is open in $\mathfrak{h}_0$, it follows that $\delta(D)\hat{} = \delta(\hat{D})$.

Now let D be an arbitrary element of $\mathfrak{J}(\mathfrak{g})$. Since π^2 is invariant under W, by Chevalley's Theorem [2(a), Lemma 9] there exists an element $\eta \in J(\mathfrak{g})$ such that $\eta(H) = \pi(H)^2$ $(H \in \mathfrak{h})$. Moreover in view of Corollary 2 of Lemma

2, we can choose an integer $m \geqq 0$ such that $\delta(D) \circ \pi^{2m}$ lies in $\mathfrak{D}(\mathfrak{h})$. Put $D_1 = D \circ \eta^m$. Then

$$\delta(D_1) = \delta(D) \circ \delta(\eta^m) = \delta(D) \circ \pi^{2m} \in \mathfrak{D}(\mathfrak{h})$$

and therefore $\delta(\hat{D}_1) = \delta(D_1)\hat{}$ by the proof above. Let q be the degree of π. Then η is homogeneous of degree $2q$. Therefore it is obvious from Lemma 12 that $\hat{\eta} = (-1)^q \partial(\eta)$ in $\mathfrak{D}(\mathfrak{g})$. Hence $\hat{D}_1 = (-1)^{qm} \hat{D} \circ \partial(\eta)^m$ and therefore

$$\begin{aligned} \delta(\hat{D}_1) &= (-1)^{qm} \delta(\hat{D}) \circ \delta(\partial(\eta^m)) \\ &= (-1)^{qm} \delta(\hat{D}) \circ \partial(\pi^{2m}) \end{aligned}$$

from [2(a), Theorem 1]. Now suppose

$$\delta(\hat{D}) = \sum_{1 \leqq i \leqq r} q_i \partial(p_i)$$

where $q_i \in Q(\mathfrak{h})$, $p_i \in S(\mathfrak{h})$ and $p_1, \cdots, p_r$ are linearly independent over $\boldsymbol{C}$. Then

$$\sum_i q_i \partial(p_i \pi^{2m}) = (-1)^{qm} \delta(D_1)\hat{} \in \mathfrak{D}(\mathfrak{h}),$$

and we conclude easily from Corollary 4 of Lemma 4 that $q_i \in P(\mathfrak{h})$. Hence $\delta(\hat{D}) \in \mathfrak{D}(\mathfrak{h})$. But D was an arbitrary elements of $\mathfrak{Z}(\mathfrak{g})$ and $D \rightarrow \hat{D}$ is an automorphism of $\mathfrak{Z}(\mathfrak{g})$. Therefore $\delta(D) \in \mathfrak{Z}(\mathfrak{h})$ for every $D \in \mathfrak{Z}(\mathfrak{g})$ and it follows from our earlier proof that $\delta(\hat{D}) = \delta(D)\hat{}$.

It remains to show that $\delta(D^*) = \delta(D)^*$ for $D \in \mathfrak{Z}(\mathfrak{g})$. Fix H_0 and U and for any $\beta \in C_c^\infty(U)$ define g_β as above. Let $\mathscr{C}_{G_0}(\mathfrak{g}_0)$ be the set of all functions in $\mathscr{C}(\mathfrak{g}_0)$ which are invariant under G_0. Then by Theorem 4 of [2(a)] there exists a unique element $f_\beta \in \mathscr{C}_{G_0}(\mathfrak{g}_0)$ such that $f_\beta = g_\beta$ on $\mathfrak{h}_0$. Then $\phi_{f_\beta} = \pi g_\beta = \sum_{s \in W} \epsilon(s) \beta^s$. But if $F_1, F_2 \in \mathscr{C}_{G_0}(\mathfrak{g}_0)$, it is clear that

$$\begin{aligned} \int F_1 F_2 \, dX &= \int |\pi(H)|^2 F_1(H) F_2(H) \, dH \\ &= (-1)^q \int \phi_{F_1} \phi_{F_2} \, dH \end{aligned}$$

where q is the degree of π. Hence if $\alpha, \beta \in C_c^\infty(U)$,

$$\int D f_\alpha \cdot f_\beta \, dX = (-1)^q [W] \int \phi_{D f_\alpha} \cdot \beta \, dH$$

since $(\phi_{D f_\alpha})^s = \epsilon(s) \phi_{D f_\alpha}$ $(s \in W)$. But $\phi_{D f_\alpha} = \delta(D) \phi_{f_\alpha}$ and $\phi_{f_\alpha} = \alpha$ on U. Hence

$$\int D f_\alpha \cdot f_\beta \, dX = (-1)^q [W] \int \delta(D) \alpha \cdot \beta \, dH.$$

On the other hand D^* is also in $\mathfrak{Z}(\mathfrak{g})$ and therefore

$$\int D f_\alpha \cdot f_\beta \, dX = \int f_\alpha \cdot D^* f_\beta \, dX = (-1)^q [W] \int \alpha \cdot \delta(D^*) \beta \, dH.$$

This proves that $(\delta(D)^* - \delta(D^*))_H = 0$ for all $H \in U$. But since U is open in $\mathfrak{h}_0$, this implies that $\delta(D)^* = \delta(D^*)$. The proof of Lemma 13 and therefore also of Theorem 1, is now quite complete.

7. Application to invariant distributions.

We now return to the notation of § 2 and take $k = \boldsymbol{R}$. Let G be a connected Lie group with Lie algebra $\mathfrak{g}$. Fix a point $Z_0 \in \mathfrak{z}'$ and an open neighborhood Ω of Z_0 in $\mathfrak{g}$.

LEMMA 14. *Let f be a locally invariant C^∞ function on Ω. Then*[3]

$$f(Z_0; D) = f(Z_0; \delta_{\mathfrak{g}/\mathfrak{z}}'(D))$$

for $D \in \mathfrak{D}(\mathfrak{g}_c)$.

This is proved in the same way as Lemma 10 of [2(e)].

LEMMA 15. *Let π denote the mapping $(x, Z) \rightarrow xZ$ of $G \times \mathfrak{z}$ into $\mathfrak{g}$. Then* rank $d\pi = n$ *everywhere on $G \times \mathfrak{z}'$. Here $n =$* dim $\mathfrak{g}$.

Since $\zeta(Z) \neq 0$, it is clear that $[Z, \mathfrak{q}] = \mathfrak{q}$ for $Z \in \mathfrak{z}'$. Our assertion can now be proved in the same way as Lemma 12 of [2(e)].

Let dx denote the Haar measure on G and dX and dZ the Euclidean measures on $\mathfrak{g}$ and $\mathfrak{z}$ respectively.

LEMMA 16. *Put $\Omega = \pi(G \times \mathfrak{z}') = G\mathfrak{z}'$. Then Ω is open in $\mathfrak{g}$ and for every $\alpha \in C_c^\infty(G \times \mathfrak{z}')$, there exists a unique element $f_\alpha \in C_c^\infty(\Omega)$ such that*

$$\int_{G \times \mathfrak{z}'} F(xZ)\alpha(x:Z)\,dx\,dZ = \int_\Omega F(X) f_\alpha(X)\,dX$$

for every $F \in C_c^\infty(\Omega)$. The mapping $\alpha \rightarrow f_\alpha$ of $C_c^\infty(G \times \mathfrak{z}')$ into $C_c^\infty(\Omega)$ is continuous and surjective and Supp $f_\alpha \subset \pi($Supp $\alpha)$.

The proof of this is similar to that of Lemma 13 of [2(e)].

COROLLARY 1. *Let F be a locally summable function on Ω. Then the function $(x, Z) \rightarrow F(xZ)$ is locally summable on $G \times \mathfrak{z}'$ and*

$$\int_{G \times \mathfrak{z}'} F(xZ)\alpha(x:Z)\,dx\,dZ = \int_\Omega F(X) f_\alpha(X)\,dX$$

for every $\alpha \in C_c^\infty(G \times \mathfrak{z}')$.

This follows from Corollary 2 of Theorem 1 of [2(e)].

---

[3] $\mathfrak{g}_c$ denotes the complexification of $\mathfrak{g}$. Similarly in other cases. For convenience we write $\delta_{\mathfrak{g}/\mathfrak{z}}'$ instead of $\delta_{\mathfrak{g}_c/\mathfrak{z}_c}'$.

LEMMA 17. *Let G_0 and $\mathfrak{z}_0$ be open subsets of G and $\mathfrak{z}'$ respectively. We assume that G_0 is connected and not empty. Let T be a locally invariant distribution on $\Omega_0 = G_0\mathfrak{z}_0$. Then there exists a unique distribution σ_T on $\mathfrak{z}_0$ such that*

$$T(f_\alpha) = \sigma_T(\beta_\alpha) \qquad (\alpha \in C_c^\infty(G_0 \times \mathfrak{z}_0))$$

where

$$\beta_\alpha(Z) = \int \alpha(x:Z)\,dx \qquad (Z \in \mathfrak{z}_0).$$

Moreover σ_T is locally invariant (with respect to $\mathfrak{z}$) and $\sigma_T = 0$ implies that $T = 0$.

Everything except the local invariance of σ_T is proved in the same way as Theorem 2 of [2(e)]. Now fix $\beta \in C_c^\infty(\mathfrak{z}_0)$ and $Y \in \mathfrak{z}$. We have to prove that $\sigma_T(\tau(Y)\beta) = 0$. Select $\gamma \in C_c^\infty(G_0)$ such that $\int \gamma(x)\,dx = 1$ and put $\alpha = \gamma \times \beta$ and $\alpha_1 = \gamma \times \tau(Y)\beta$. Then if $F \in C_c^\infty(\Omega_0)$, we have

$$\int f_{\alpha_1}(X)F(X)\,dX = \int \gamma(x)\beta(Z;\tau(Y))F(xZ)\,dx dZ.$$

But it is clear that

$$\int \beta(Z;\tau(Y))F(xZ)\,dx dZ = \int \beta(Z)F(x;Y:Z)\,dx dZ$$

where $F(x:Z) = F(xZ)$ $(x \in G_0, Z \in \mathfrak{z}_0)$. Therefore

$$\int f_{\alpha_1}F\,dX = -\int \gamma(x;Y)\beta(Z)F(x:Z)\,dx dZ = -\int f_{\alpha_2}F\,dX$$

where $\alpha_2 = Y\gamma \times \beta$. This proves that $f_{\alpha_1} = -f_{\alpha_2}$ and therefore

$$\sigma_T(\tau(Y)\beta) = T(f_{\alpha_1}) = -T(f_{\alpha_2}) = -\sigma_T(\beta_{\alpha_2}).$$

But since

$$\int \gamma(x;Y)\,dx = 0,$$

it is clear that $\beta_{\alpha_2} = 0$ and therefore $\sigma_T(\tau(Y)\beta) = 0$.

COROLLARY 1. *Assume that $G_0 = G$ and T is invariant under G. Let Ξ be the subgroup of all $\xi \in G$ such that $\xi\mathfrak{z}_0 = \mathfrak{z}_0$. Then σ_T is invariant under Ξ.*

Let Ξ_1 be the normalizer of $\mathfrak{z}$ in G and Ξ_0 the analytic subgroup of G corresponding to $\mathfrak{z}$. Then Ξ_0 is the connected component of 1 in Ξ_1 and if Z_G is the center of G, $\Xi_1/Z_G\Xi_0$ is finite (see [2(b), Lemma 15]). Since $\mathfrak{z}$ is reductive, the measure dZ is invariant under Ξ_0 and therefore also under Ξ_1. Now we may obviously assume that $\mathfrak{z}_0$ is not empty. Since it is open in $\mathfrak{z}$, it spans $\mathfrak{z}$ over $\boldsymbol{R}$. Hence $\Xi \subset \Xi_1$ and this proves that dZ is invariant under Ξ. Now fix γ as above and for given $\beta \in C_c^\infty(\mathfrak{z}_0)$ and $\xi \in \Xi$, define $\beta_1(Z)$

$=\beta(\xi Z)$ $(Z \in \mathfrak{z}_0)$. Then $\sigma_T(\beta_1) = T(f_{\alpha_1})$ where $\alpha_1 = \gamma \times \beta_1$. On the other hand

$$\begin{aligned} \int f_{\alpha_1} F\, dX &= \int \gamma(x)\beta(\xi Z) F(xZ)\, dx dZ \\ &= \int \gamma(x\xi)\beta(Z) F(xZ)\, dx dZ \end{aligned}$$

for any $F \in C_c^\infty(\Omega_0)$. Put $\alpha(x:Z) = \gamma(x\xi)\beta(Z)$ $(x \in G, Z \in \mathfrak{z}_0)$. Then the above equation implies that $f_{\alpha_1} = f_\alpha$ and therefore

$$\sigma_T(\beta_1) = T(f_{\alpha_1}) = T(f_\alpha) = \sigma_T(\beta),$$

since

$$\int \alpha(x:Z)\, dx = \beta(Z) \qquad (Z \in \mathfrak{z}_0).$$

This proves the corollary.

COROLLARY 2.[4] *For any $D \in \mathfrak{J}(\mathfrak{g}_c)$, write $\Delta(D) = \delta_{\mathfrak{g}/\mathfrak{z}}'(D)$. Then $\sigma_{DT} = \Delta(D)\sigma_T$ under the conditions of Lemma* 17.

Put $T' = DT$. Then T' is also locally invariant and $T'(f_\alpha) = T(D^* f_\alpha)$ for $\alpha \in C_c^\infty(G_0 \times \mathfrak{z}_0)$. Fix $F \in C_c^\infty(\Omega_0)$ and put $F_1 = DF$. Then

$$\begin{aligned} \int f_\alpha(X; D^*) F(X)\, dX &= \int f_\alpha F_1\, dX \\ &= \int \alpha(x:Z) F_1(xZ)\, dx dZ. \end{aligned}$$

As usual, define $F(x:Z) = F(xZ)$ $(x \in G_0, Z_0 \in \mathfrak{z}_0)$. Then since $D \in \mathfrak{J}(\mathfrak{g}_c)$, it is clear that

$$F_1(xZ) = F(xZ; D) = F(xZ; D^x).$$

Moreover from Corollary 2 of Lemma 2, we can choose $g_i \in \mathfrak{S}_+(\mathfrak{q}_c)$, $p_i \in S(\mathfrak{z}_c)$ and $a_i \in C^\infty(\mathfrak{z}_0)$ $(1 \leqq i \leqq r)$ such that

$$D_Z = (\Delta(D))_Z + \sum_i a_i(Z)\partial(\Gamma_Z(g_i \otimes p_i))$$

for $Z \in \mathfrak{z}_0$. Since

$$(D^x)_{xZ} = (D_Z)^x$$

it follows that

$$F(xZ; D^x) = F(x:Z; \Delta(D)) + \sum_i F(x; g_i : Z; \xi_i)$$

for $x \in G_0$ and $Z \in \mathfrak{z}_0$. Here ξ_i is the differential operator $a_i\partial(p_i)$ on $\mathfrak{z}_0$. Put $\alpha_0(x:Z) = \alpha(x:Z; \Delta(D)^*)$ and $\alpha_i(x:Z) = \alpha(x; g_i^* : Z; \xi_i^*)$ $(1 \leqq i \leqq r)$, where the star denotes the adjoint, as usual. Then

$$\int D^* f_\alpha \cdot F\, dX = \sum_{0 \leqq i \leqq r} \int \alpha_i(x:Z) F(xZ)\, dx dZ$$

---

[4] This can be regarded as a generalization of Lemma 14 to distributions.

$$= \sum_{0\leqq i\leqq r} \int f_{\alpha_i} F\, dX.$$

This proves that

$$D^* f_\alpha = \sum_{0\leqq i\leqq r} f_{\alpha_i}$$

and therefore

$$T'(f_\alpha) = \sum_{0\leqq i\leqq r} \sigma_T(\beta_{\alpha_i}).$$

But since $g_i \in \mathfrak{S}_+(\mathfrak{q}_c)$ $(1 \leqq i \leqq r)$, it is clear that $g_i h = 0$ where h is the constant function 1 on G. Hence

$$\int g_i^* f\, dx = \int f \cdot g_i h\, dx = 0$$

for any $f \in C_c^\infty(G)$. This shows that $\beta_{\alpha_i} = 0$ for $i \geqq 1$ and therefore

$$T'(f_\alpha) = \sigma_T(\beta_{\alpha_0}) = \sigma_T(\Delta(D)^* \beta_\alpha).$$

Since the mapping $\alpha \to \beta_\alpha$ of $C_c^\infty(G_0 \times \mathfrak{z}_0)$ into $C_c^\infty(\mathfrak{z}_0)$ is obviously surjective, this proves that $\sigma_{T'} = \Delta(D)\sigma_T$.

COROLLARY 3. *Let* $D_1, D_2 \in \mathfrak{Z}(\mathfrak{g}_c)$. *Then*

$$\Delta(D_1 D_2)\sigma_T = \Delta(D_1)\Delta(D_2)\sigma_T$$

in the notation of Corollary 2 above.

This follows from the fact that $(D_1 D_2)T = D_1(D_2 T)$.

8. An explicit formula. Let $p \to p_{\mathfrak{z}}$ $(p \in S(\mathfrak{g}_c)$ denote the natural projection of $S(\mathfrak{g}_c)$ onto $S(\mathfrak{z}_c)$ with respect to the decomposition $\mathfrak{g}_c = \mathfrak{z}_c + \mathfrak{q}_c$ (see § 3). Define $\zeta \in P(\mathfrak{z}_c)$ as in § 2 and put[5]

$$\zeta^{-\frac{1}{2}} D \circ \zeta^{\frac{1}{2}} = \mu_\zeta(D) \qquad (D \in \mathfrak{D}(\mathfrak{z}_c))$$

in the notation of Lemma 8.

Let $I(\mathfrak{g}_c)$ be the algebra of all invariants of G in $S(\mathfrak{g}_c)$. We keep to the notation of § 7.

THEOREM 2. *Let T be a locally invariant distribution on* Ω_0. *Then*

$$\Delta(\partial(p))\sigma_T = (\zeta^{-\frac{1}{2}} \partial(p_{\mathfrak{z}}) \circ \zeta^{\frac{1}{2}})\sigma_T$$

for $p \in I(\mathfrak{g}_c)$.

Let $\mathfrak{c}$ be the center and $\mathfrak{g}_1$ the derived algebra of $\mathfrak{g}$. We can obviously choose a real, symmetric and nondegenerate bilinear form B on $\mathfrak{g}$ such that

---

[5] Actually it is not difficult to show that there exists an invariant polynomial function η on $\mathfrak{z}$ such that $\eta^2 = \zeta$.

1) $B(X,Y)=\mathrm{tr}(\mathrm{ad}\,X\,\mathrm{ad}\,Y)$ for $X,Y\in\mathfrak{g}_1$ and 2) $B(X,Y)=0$ for $X\in\mathfrak{c}$ and $Y\in\mathfrak{g}_1$. Then B is invariant under G. Extend this form on $\mathfrak{g}_c$ by linearity over $\boldsymbol{C}$ and identify $\mathfrak{g}_c$ with its dual under B. Then $S(\mathfrak{g}_c)=P(\mathfrak{g}_c)$ and it follows from Lemma 1 that $B_{\mathfrak{z}}$, the restriction of B on $\mathfrak{z}_c$, is also nondegenerate. We identify $\mathfrak{z}_c$ with its dual under $B_{\mathfrak{z}}$. Then it is easy to verify that $p_{\mathfrak{z}}$ is simply the restriction of the polynomial function p on $\mathfrak{z}_c$ ($p\in S(\mathfrak{g}_c)$).

Let ω denote the polynomial function $X\to B(X,X)$ on $\mathfrak{g}_c$. We shall first prove the following result.

LEMMA 18. $\Delta(\partial(\omega)=\zeta^{-\frac{1}{2}}\partial(\omega_{\mathfrak{z}})\circ\zeta^{\frac{1}{2}}$.

Fix a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{z}$ and let $\mathfrak{h}'$ be the set of all elements of $\mathfrak{h}$ which are regular in $\mathfrak{g}$. Put

$$D=\Delta(\partial(\omega))-\zeta^{-\frac{1}{2}}\partial(\omega_{\mathfrak{z}})\circ\zeta^{\frac{1}{2}}.$$

Let Ξ be the analytic subgroup of G corresponding to $\mathfrak{z}$. It follows from Lemma 10 that D is invariant under Ξ. Moreover $\mathfrak{z}_0'=\bigcup_{\xi\in\Xi}\xi\mathfrak{h}'$ is an open and non-empty subset of $\mathfrak{z}'$ (Lemma 15). Therefore since $D\in\mathfrak{D}_r(\mathfrak{z}_0)$, it would be sufficient to verify that $D_{\xi H}=0$ for $\xi\in\Xi$ and $H\in\mathfrak{h}'$. But $D_{\xi H}=(D_H)^{\xi}$ because D is invariant under Ξ. Hence we have only to prove that $D_H=0$ for $H\in\mathfrak{h}'$.

Let P be the set of all positive roots of $(\mathfrak{g}_c,\mathfrak{h}_c)$ (under some fixed order). For any root α define X_α as in the proof of Lemma 1. We normalize X_α and $X_{-\alpha}$ in such a way that $B(X_\alpha,X_{-\alpha})=1$. The restriction $B_{\mathfrak{h}}$ of B on $\mathfrak{h}_c$ is also nondegenerate. We identify $\mathfrak{h}_c$ with it dual under $B_{\mathfrak{h}}$.

Define subsets $P_{\mathfrak{z}}$ and $P_{\mathfrak{q}}$ of P as follows. A root $\alpha\in P$ belongs to $P_{\mathfrak{z}}$ or $P_{\mathfrak{q}}$ according as X_α lies in $\mathfrak{z}_c$ or $\mathfrak{q}_c$. Then (see [2(a), p. 98])

$$\omega=\omega_{\mathfrak{z}}+2\sum_{\alpha\in P_{\mathfrak{q}}}X_\alpha X_{-\alpha},$$

$$\omega_{\mathfrak{z}}=\omega_{\mathfrak{h}}+2\sum_{\alpha\in P_{\mathfrak{z}}}X_\alpha X_{-\alpha},$$

where $\omega_{\mathfrak{h}}$ is the restriction of ω on $\mathfrak{h}_c$. A simple computation shows that

$$\Gamma_H(\lambda(X_\alpha X_{-\alpha})\otimes 1)=-\alpha(H)^2X_\alpha X_{-\alpha}+\alpha(H)H_\alpha \qquad (\alpha\in P, H\in\mathfrak{h})$$

where λ is the canonical mapping of $S(\mathfrak{g}_c)$ onto $\mathfrak{G}$. Here H_α is the element in $\mathfrak{h}_c$ which corresponds to α under $B_{\mathfrak{h}}$. Therefore

$$\omega=\omega_{\mathfrak{z}}+2\sum_{\alpha\in P_{\mathfrak{q}}}\alpha(H)^{-1}H_\alpha-2\sum_{\alpha\in P_{\mathfrak{q}}}\alpha(H)^{-2}\Gamma_H(\lambda(X_\alpha X_{-\alpha})\otimes 1)$$

for $H\in\mathfrak{h}'$. This implies that

$$(\Delta(\partial(\omega)))_H = \partial(\omega_{\mathfrak{z}}) + 2 \sum_{\alpha \in P_{\mathfrak{q}}} \alpha(H)^{-1} \partial(H_\alpha) \qquad (H \in \mathfrak{h}').$$

We find in the same way that

$$\omega_{\mathfrak{z}} = \omega_{\mathfrak{h}} + 2 \sum_{\alpha \in P_{\mathfrak{z}}} \alpha(H)^{-1} H_\alpha - 2 \sum_{\alpha \in P_{\mathfrak{z}}} \alpha(H)^{-2} \Gamma_H(\lambda(X_\alpha X_{-\alpha}) \otimes 1) \qquad (H \in \mathfrak{h}').$$

Now fix a point $H_0 \in \mathfrak{h}'$. Choose an open convex neighborhood U_c of H_0 in $\mathfrak{z}_c$ such that 1) every point of U_c is regular in $\mathfrak{g}_c$ and 2) there exists a holomorphic function η on U_c such that $\zeta = \eta^2$ on U_c. Put $U = U_c \cap \mathfrak{z}$. Then if ϕ is a C^∞ function on U, it follows from Lemma 3 of [2(a)] that

$$\phi(H; \partial(\omega_{\mathfrak{z}})) = \phi(H; \partial(\omega_{\mathfrak{h}})) + 2 \sum_{\alpha \in P_{\mathfrak{z}}} \alpha(H)^{-1} \phi(H; \partial(H_\alpha))$$
$$- 2 \sum_{\alpha \in P_{\mathfrak{z}}} \alpha(H)^{-2} \phi(1; \lambda(X_\alpha X_{-\alpha}) : H) \qquad (H \in U \cap \mathfrak{h}).$$

Now fix $f \in C^\infty(U)$. It is enough to show that $f(H_0; D) = 0$. Put $g = \eta f$. Then

$$f(H; D) = f(H; \Delta(\partial(\omega))) - \eta(H)^{-1} g(H; \partial(\omega_{\mathfrak{z}})) \qquad (H \in U \cap \mathfrak{h}).$$

Put

$$\pi = \prod_{\alpha \in P} \alpha, \quad \pi_{\mathfrak{q}} = \prod_{\alpha \in P_{\mathfrak{q}}} \alpha \text{ and } \pi_{\mathfrak{z}} = \prod_{\alpha \in P_{\mathfrak{z}}} \alpha.$$

It is obvious that $\zeta(H) = (-1)^r \pi_{\mathfrak{q}}(H)^2$ for $H \in \mathfrak{h}_c$. (Here r is the number of roots in $P_{\mathfrak{q}}$.) Since the set $\mathfrak{h}_c \cap U_c$ is convex and therefore connected, it is clear that $\eta(H) = \epsilon \pi_{\mathfrak{q}}(H)$ $(H \in \mathfrak{h}_c \cap U_c)$ where ϵ is a constant such that $\epsilon^4 = 1$.

Now $g(1; \lambda(X_\alpha X_{-\alpha}) : H) = g(H; \tau(\lambda(X_\alpha X_{-\alpha})))$ for $\alpha \in P_{\mathfrak{z}}$ (see Lemma 11 of [2(e)]). Since ζ and therefore also η are locally invariant (with respect to $\mathfrak{z}$), it is clear that

$$\tau(\lambda(X_\alpha X_{-\alpha})) g = \eta \tau(\lambda(X_\alpha X_{-\alpha})) f.$$

Hence we conclude from our formula above that

$$g(H; \delta(\omega_{\mathfrak{z}})) = \epsilon f(H; \partial(\omega_{\mathfrak{h}}) \circ \pi_{\mathfrak{q}}) + 2 \sum_{\alpha \in P_{\mathfrak{z}}} \alpha(H)^{-1} \epsilon f(H; \partial(H_\alpha) \circ \pi_{\mathfrak{q}})$$
$$- 2\epsilon \sum_{\alpha \in P_{\mathfrak{z}}} \alpha(H)^{-2} \pi_{\mathfrak{q}}(H) f(1; \lambda(X_\alpha X_{-\alpha}) : H)$$

for $H \in U \cap \mathfrak{h}$. On the other hand, it follows from our formula for $(\Delta(\partial(\omega)))_H$ that

$$f(H; \Delta(\partial(\omega))) = f(H; \partial(\omega_{\mathfrak{z}})) + 2 \sum_{\alpha \in P_{\mathfrak{q}}} \alpha(H)^{-1} f(H; \partial(H_\alpha))$$

$$= f(H;\partial(\omega_{\mathfrak{h}})) + 2\sum_{\alpha\in P}\alpha(H)^{-1}f(H;\partial(H_\alpha))$$
$$- 2\sum_{\alpha\in P_{\mathfrak{z}}}\alpha(H)^{-2}f(1;\lambda(X_\alpha X_{-\alpha}):H)$$

This proves that

$$\pi_{\mathfrak{q}}(H)f(H;D) = -f(H;\Box) \qquad (H\in U\cap\mathfrak{h})$$

where $\Box$ is the element of $\mathfrak{D}_r(\mathfrak{h}_c)$ given by[6]

$$\Box = \{\partial(\omega_{\mathfrak{h}}),\pi_{\mathfrak{q}}\} + 2\sum_{\alpha\in P_{\mathfrak{z}}}\alpha^{-1}(\partial(H_\alpha)\pi_{\mathfrak{q}}) - 2\pi_{\mathfrak{q}}\sum_{\alpha\in P_{\mathfrak{q}}}\alpha^{-1}\partial(H_\alpha).$$

We shall now verify that $\Box = 0$ and this would prove the required result.

Let $\alpha_1,\alpha_2,\cdots,\alpha_r$ be all the distinct roots in $P_{\mathfrak{q}}$. Then $\pi_{\mathfrak{q}} = \alpha_1\alpha_2\cdots\alpha_r$ and $\{\partial(\omega_{\mathfrak{h}}),\alpha_i\} = 2\partial(\alpha_i)$ (see [2(a), p. 98]). Hence

$$\{\partial(\omega_{\mathfrak{h}}),\pi_{\mathfrak{q}}\} = 2\sum_{1\leqq i\leqq r}\alpha_1\cdots\alpha_{i-1}\partial(\alpha_i)\circ(\alpha_{i+1}\cdots\alpha_r)$$
$$= 2\pi_{\mathfrak{q}}\sum_{1\leqq i\leqq r}\alpha_i^{-1}\partial(\alpha_i) + q$$

where $q\in S(\mathfrak{h}_c)$. Hence

$$\Box = q + 2\sum_{\alpha\in P_{\mathfrak{z}}}\alpha^{-1}(\partial(H_\alpha)\pi_{\mathfrak{q}}) \in Q(\mathfrak{h}_c).$$

Therefore applying the differential operator $\Box$ to the constant function 1, we get

$$\Box = \Box 1 = \partial(\omega_{\mathfrak{h}})\pi_{\mathfrak{q}} + 2\sum_{\alpha\in P_{\mathfrak{z}}}\alpha^{-1}(\partial(H_\alpha)\pi_{\mathfrak{q}}).$$

On the other hand if f_0 denotes the constant function 1 on U, we have

$$\pi_{\mathfrak{q}}(H)f_0(H;D) = -f_0(H;\Box) = -\Box(H) \qquad (H\in U\cap\mathfrak{h}).$$

But it follows from our formula for $(\Delta(\partial(\omega)))_H$ that

$$f_0(H;D) = -\eta(H)^{-1}\eta(H;\partial(\omega_{\mathfrak{z}})).$$

Therefore

$$\epsilon\Box(H) = \eta(H;\partial(\omega_{\mathfrak{z}})) \qquad (H\in U\cap\mathfrak{h}).$$

But since η is locally invariant (with respect to $\mathfrak{z}$), we conclude from Lemma 14 that

$$\eta(H;\partial(\omega_{\mathfrak{z}})) = \eta(H;\delta_{\mathfrak{z}/\mathfrak{h}}'(\partial(\omega_{\mathfrak{z}}))).$$

---

[6] If D_1, D_2 are two differential operators, we denote $D_1\circ D_2 - D_2\circ D_1$, as usual, by $\{D_1,D_2\}$.

But $\delta_{\mathfrak{z}/\mathfrak{h}}{}' \ (\partial(\omega_{\mathfrak{z}})) = \pi_{\mathfrak{z}}^{-1}\partial(\omega_{\mathfrak{h}}) \circ \pi_{\mathfrak{z}}$ from [2(a), Lemma 8]. Hence

$$\epsilon\Box(H) = \eta(H; \pi_{\mathfrak{z}}^{-1}\, \partial(\omega_{\mathfrak{h}}) \circ \pi_{\mathfrak{z}}) = \epsilon\pi_{\mathfrak{z}}(H)^{-1}\pi(H; \partial(\omega_{\mathfrak{h}})).$$

However we know that $\partial(\omega_{\mathfrak{h}})\pi = 0$ [2(a), p. 101]. Therefore $\Box = 0$ and so the lemma is proved.

Now we come to the proof of Theorem 2. It is obviously enough to consider the case when p is homogeneous. Let Λ denote the derivation of $\mathfrak{J}(\mathfrak{g}_c)$ given by $\Lambda D = (1/2)\{\partial(\omega), D\}$ $(D \in \mathfrak{J}(\mathfrak{g}_c))$. Similarly let ν_0 denote the derivation of $\mathfrak{D}_r(\mathfrak{z}_c)$ defined by $\nu_0\xi = (1/2)\{\Delta(\partial(\omega)), \xi\}$ $(\xi \in \mathfrak{D}_r(\mathfrak{z}_c))$. Then it follows from Corollary 3 of Lemma 17 that

$$\Delta(\Lambda D)\sigma_T = (\nu_0\Delta(D))\sigma_T \qquad (D \in \mathfrak{J}(\mathfrak{g}_c))$$

for a locally invariant distribution T on Ω_0. Hence we conclude by induction on m, that

$$\Delta(\Lambda^m D)\sigma_T = (\nu_0{}^m\Delta(D))\sigma_T \qquad (m \geqq 0).$$

Let p be a homogeneous element in $I(\mathfrak{g}_c)$ of degree k. It is clear from its definition that $\Delta(p) = p_{\mathfrak{z}}$. Moreover $\Lambda^k p = k!\partial(p)$ (see [2(a), p. 99]). Therefore

$$k!\Delta(\partial(p))\sigma_T = (\nu_0{}^k p_{\mathfrak{z}})\sigma_T.$$

But $\Delta(\partial(\omega)) = \mu_\zeta(\partial(\omega_{\mathfrak{z}}))$ from Lemma 18 and $p_{\mathfrak{z}} = \mu_\zeta(p_{\mathfrak{z}})$. Therefore since μ_ζ is an automorphism of $\mathfrak{D}_r(\mathfrak{z}_c)$, it is clear that

$$\nu_0{}^k p_{\mathfrak{z}} = \mu_\zeta(\nu^k p_{\mathfrak{z}})$$

where ν is the derivation of $\mathfrak{D}_r(\mathfrak{z}_c)$ given by

$$\nu\xi = (1/2)\{\partial(\omega_{\mathfrak{z}}), \xi\} \qquad (\xi \in \mathfrak{D}_r(\mathfrak{z}_c)).$$

Moreover $p_{\mathfrak{z}}$ is also a homogeneous element of $S(\mathfrak{z}_c)$ of degree k and $\nu^k p_{\mathfrak{z}} = k!\partial(p_{\mathfrak{z}})$ [2(a), p. 99]. Therefore

$$\mu_\zeta(\nu^k p_{\mathfrak{z}}) = k!\mu_\zeta(\partial(p_{\mathfrak{z}}))$$

and this proves that

$$\Delta(\partial(p))\sigma_T = \mu_\zeta(\partial(p_{\mathfrak{z}}))\sigma_T.$$

9. **An elementary result.** For $X \in \mathfrak{g}_c$, define the derivation d_X of $S(\mathfrak{g}_c)$ as in [2(e), § 7]. Then it is obvious that $I(\mathfrak{g}_c)$ consists of all $p \in S(\mathfrak{g}_c)$ such that $d_X p = 0$ for every $X \in \mathfrak{g}_c$. Since $[\mathfrak{z}, \mathfrak{q}] \subset \mathfrak{q}$, it follows that if $p \in I(\mathfrak{g}_c)$ then $d_Z p_{\mathfrak{z}} = 0$ $(Z \in \mathfrak{z}_c)$ and therefore $p_{\mathfrak{z}} \in I(\mathfrak{z}_c)$. Fix a Cartan

subalgebra $\mathfrak{h}$ of $\mathfrak{z}$ and let $W(\mathfrak{g})$ and $W(\mathfrak{z})$ denote the Weyl groups of $(\mathfrak{g}_c, \mathfrak{h}_c)$ and $(\mathfrak{z}_c, \mathfrak{h}_c)$ respectively. Then $W(\mathfrak{z})$ is generated by the Weyl reflexions s_α corresponding to the roots α in $P_{\mathfrak{z}}$ (in the notation of the proof of Lemma 18). Hence we can regard $W(\mathfrak{z})$ as a subgroup of $W(\mathfrak{g})$. Let $w(\mathfrak{g}/\mathfrak{z})$ denote the index $[W(\mathfrak{g}) : W(\mathfrak{z})]$.

LEMMA 19. *Let J denote the image of $I(\mathfrak{g}_c)$ in $I(\mathfrak{z}_c)$ under the mapping $p \rightarrow p_{\mathfrak{z}}$ $(p \in I(\mathfrak{g}_c))$. Then $I(\mathfrak{z}_c)$ is a free abelian module over J of rank $w(\mathfrak{g}/\mathfrak{z})$.*

Let $\mathfrak{p}$ be the subspace of $\mathfrak{z}$ complementary to $\mathfrak{h}$ such that $[\mathfrak{h}, \mathfrak{p}] \subset \mathfrak{p}$ (see Lemma 1). Then we have the decompositions $\mathfrak{z} = \mathfrak{h} + \mathfrak{p}$ and $\mathfrak{g} = \mathfrak{h} + (\mathfrak{p} + \mathfrak{q})$. Let $E_{\mathfrak{z}}$ and $E_{\mathfrak{g}}$ denote the corresponding projections of $S(\mathfrak{z}_c)$ and $S(\mathfrak{g}_c)$ respectively on $S(\mathfrak{h}_c)$ (see § 3). Define $p_{\mathfrak{h}} = E_{\mathfrak{g}} p$ $(p \in I(\mathfrak{g}_c))$, $q_{\mathfrak{h}} = E_{\mathfrak{z}} q$ $(q \in I(\mathfrak{z}_c))$, $J_{\mathfrak{g}} = E_{\mathfrak{g}} I(\mathfrak{g}_c)$, $J_{\mathfrak{z}} = E_{\mathfrak{z}} I(\mathfrak{z}_c)$. Then by Chevalley's Theorem [2(a), Lemma 9], $J_{\mathfrak{z}}$ is the algebra of all invariants of $W(\mathfrak{z})$ in $S(\mathfrak{h}_c)$ and $E_{\mathfrak{z}}$ defines an isomorphism of $I(\mathfrak{z}_c)$ onto $J_{\mathfrak{z}}$. Similarly $E_{\mathfrak{g}}$ defines an isomorphism of $I(\mathfrak{g}_c)$ onto $J_{\mathfrak{g}}$ which is, in fact, the algebra of all invariants of $W(\mathfrak{g})$ in $S(\mathfrak{h}_c)$. Since $(p_{\mathfrak{z}})_{\mathfrak{h}} = p_{\mathfrak{h}}$ for $p \in I(\mathfrak{g}_c)$, it is clear that under the above isomorphism of $I(\mathfrak{z}_c)$ onto $J_{\mathfrak{z}}$, J is mapped onto $J_{\mathfrak{g}}$. Therefore it would be sufficient to show that $J_{\mathfrak{z}}$ is a free abelian module over $J_{\mathfrak{g}}$ of rank $w(\mathfrak{g}/\mathfrak{z})$. The proof of this is substantially the same as that of Lemma 8 of [2(c)].

10. An example.

As an application of the above theory, we shall conclude this paper with a proof of the following theorem.

THEOREM 3. *Let $\mathfrak{g}$ be a reductive Lie algebra over $\boldsymbol{R}$ and S the set of all singular elements of $\mathfrak{g}$. Let T be a distribution on an open subset Ω of $\mathfrak{g}$. Assume that:*

1) *T is locally invariant;*
2) *There exists an ideal $\mathfrak{U}$ in $I(\mathfrak{g}_c)$ such that $\dim I(\mathfrak{g}_c)/\mathfrak{U} < \infty$ and $\partial(u) T = 0$ for $u \in \mathfrak{U}$;*
3) $\operatorname{Supp} T \subset \Omega \cap S$.

Then $T = 0$.

We use induction on $\dim \mathfrak{g}$. Let us assume that contrary to our assertion, $T \neq 0$. Consider the space $\mathfrak{T}$ of all distributions on Ω of the form $\partial(p) T = 0$ $(p \in I(\mathfrak{g}_c))$. Then it is obvious from condition 2) that $\dim \mathfrak{T} < \infty$. It is also obvious that every distribution in $\mathfrak{T}$ satisfies all the conditions of the theorem and $\mathfrak{T} \neq \{0\}$ since $T \in \mathfrak{T}$.

For any p in $I(\mathfrak{g}_c)$, let $\rho(p)$ denote the linear transformation $T' \to \partial(p)T'$ $(T' \in \mathfrak{T})$ in $\mathfrak{T}$. Then ρ is a representation of the abelian algebra $I(\mathfrak{g}_c)$ on the finite-dimensional space $\mathfrak{T} \neq \{0\}$. Hence we can choose an element $T_0 \neq 0$ in $\mathfrak{T}$ and a homomorphism χ of $I(\mathfrak{g}_c)$ into $\boldsymbol{C}$ such that $\partial(p)T_0 = \chi(p)T_0$ for every $p \in I(\mathfrak{g}_c)$. We intend to derive a contradiction by proving that $T_0 = 0$.

Let us change notation and write T instead of T_0. Fix an element $Z_0 \in \boldsymbol{\Omega}$. We have to show that $T = 0$ around Z_0. Let $\mathfrak{c}$ be the center and $\mathfrak{g}_1$ the derived algebra of $\mathfrak{g}$ so that $\mathfrak{g}$ is the direct sum of $\mathfrak{c}$ and $\mathfrak{g}_1$. If $\mathfrak{g} = \mathfrak{c}$, it is clear that S is empty and therefore $T = 0$. Hence we may suppose that $\mathfrak{g}_1 \neq \{0\}$. First assume that $\dim \mathfrak{g}_1 < \dim \mathfrak{g}$. Choose open connected sets $\mathfrak{c}_0$ and Ω_1 in $\mathfrak{c}$ and $\mathfrak{g}_1$ respectively such that $Z_0 \in \mathfrak{c}_0 + \Omega_1 \subset \Omega$. Then it would be enough (see [2(e), Lemma 3]) to show that $T(\alpha \times \beta) = 0$ for $\alpha \in C_c^\infty(\mathfrak{c}_0)$ and $\beta \in C_c^\infty(\Omega_1)$. Since $\mathfrak{c} \subset I(\mathfrak{g}_c)$, $\partial(C)T = \chi(C)T$ for $C \in \mathfrak{c}$. Now fix β and put $T_\beta(\alpha) = T(\alpha \times \beta)$ for $\alpha \in C_c^\infty(\mathfrak{c}_0)$. Then T_β is a distribution on $\mathfrak{c}_0$. Let $C_1, \cdots, C_r$ be a base for $\mathfrak{c}$ over $\boldsymbol{R}$. Then $\partial(C_1^2 + C_2^2 + \cdots + C_r^2)$ is an elliptic differential operator on $\mathfrak{c}$. Since

$$\partial(C_1^2 + \cdots + C_r^2)T_\beta = \chi(C_1^2 + \cdots + C_r^2)T_\beta,$$

it follows that T_β is an analytic function on $\mathfrak{c}_0$. Let dC denote the Euclidean measure over $\mathfrak{c}$. Then it follows from the differential equations $\partial(C)T_\beta = \chi(C)T_\beta$ $(C \in \mathfrak{c})$, that

$$T_\beta(\alpha) = \nu(\beta) \int \alpha(C) e^{\chi(C)}\, dC \qquad (\alpha \in C_c^\infty(\mathfrak{c}_0))$$

where $\nu(\beta)$ is a complex number independent of α. Choose $\alpha_0 \in C_c^\infty(\mathfrak{c}_0)$, such that

$$\int \alpha_0(C) e^{\chi(C)}\, dC = 1.$$

Then

$$\nu(\beta) = T_\beta(\alpha_0) = T(\alpha_0 \times \beta)$$

for $\beta \in C_c^\infty(\Omega_1)$. Hence ν is a distribution on Ω_1. It is obvious that $\tau(X)\nu = 0$ for $X \in \mathfrak{g}_1$ and $\partial(p)\nu = \chi(p)\nu$ for $p \in I(\mathfrak{g}_{1c})$. Moreover if $X \in \mathfrak{g}_1$ and $C \in \mathfrak{c}$, then $X + C$ is singular in $\mathfrak{g}$ if and only if X is singular in $\mathfrak{g}_1$. Hence $\operatorname{Supp} \nu \subset \Omega_1 \cap S_1$ where $S_1 = \mathfrak{g}_1 \cap S$. Since $\dim \mathfrak{g}_1 < \dim \mathfrak{g}$, our induction hypothesis is applicable to $\mathfrak{g}_1$ and therefore $\nu = 0$. But then

$$T(\alpha \times \beta) = \nu(\beta) \int \alpha(C) e^{\chi(C)}\, dC = 0$$

for $\alpha \in C_c^\infty(\mathfrak{c}_0)$ and $\beta \in C_c^\infty(\Omega_1)$. This proves that $T = 0$ around Z_0.

So we can now assume that $\mathfrak{g}_1 = \mathfrak{g}$ and therefore $\mathfrak{g}$ is semisimple. Then $Z_0 = H_0 + X_0$ where H_0 is a semisimple and X_0 a nilpotent element of $\mathfrak{g}$ and $[H_0, X_0] = 0$ (see [1, p. 79]). First assume $H_0 \neq 0$. Let $\mathfrak{z}$ be the centralizer of H_0 in $\mathfrak{g}$. Then since H_0 is contained in some Cartan subalgebra of $\mathfrak{g}$, it is clear that $\mathfrak{z}$ is reductive in $\mathfrak{g}$ and $\operatorname{rank} \mathfrak{z} = \operatorname{rank} \mathfrak{g}$. Define $\mathfrak{q}$ as in Lemma 1. Since $X_0 \in \mathfrak{z}$, $\operatorname{ad} X_0$ maps $\mathfrak{q}$ into itself. Also the eigenvalues of $\operatorname{ad} Z_0 = \operatorname{ad}(H_0 + X_0)$ in $\mathfrak{q}$ are the same as those of $\operatorname{ad} H_0$. Hence none of them can be zero. Therefore $\zeta(Z_0) \neq 0$ and so $Z_0 \in \mathfrak{z}'$ in the notation of § 2.

Choose open connected neighborhoods G_0 and $\mathfrak{z}_0$ of 1 and Z_0 in G and $\mathfrak{z}'$ respectively such that $\Omega_0 = G_0 \mathfrak{z}_0 \subset \Omega$. We now use the notation of §§ 7-9. Then in view of Lemma 17, it would be enough to prove that $\sigma_T = 0$. Fix $\beta \in C_c^\infty(\mathfrak{z}_0)$ and assume that every element of $\operatorname{Supp} \beta$ is regular in $\mathfrak{z}$. Choose $\gamma \in C_c^\infty(G_0)$ such that $\int \gamma \, dx = 1$ and put $\alpha = \gamma \times \beta$. Then $\sigma_T(\beta) = T(f_\alpha)$. Let $X \in \operatorname{Supp} f_\alpha$. Then by Lemma 16, $X = xZ$ for some $x \in G_0$ and $Z \in \operatorname{Supp} \beta$. Therefore Z is regular in $\mathfrak{z}$ and $\zeta(Z) \neq 0$. This obviously means that Z, and hence also X, is regular in $\mathfrak{g}$. Therefore $\operatorname{Supp} f_\alpha$ does not meet S and hence $T(f_\alpha) = 0$. This proves that every point of $\operatorname{Supp} \sigma_T$ is singular in $\mathfrak{z}$. Since ζ takes only real and non-zero values on $\mathfrak{z}_0$, it is clear $|\zeta|^2 = \zeta^2$ on $\mathfrak{z}_0$ and $|\zeta|^{\frac{1}{2}}$, $|\zeta|^{-\frac{1}{2}}$ are analytic functions on $\mathfrak{z}_0$. Therefore it is clear from Lemma 8 that $\mu_\zeta(D) = |\zeta|^{-\frac{1}{2}} D \circ |\zeta|^{\frac{1}{2}}$ $(D \in \mathfrak{D}(\mathfrak{z}_c))$ as differential operators on $\mathfrak{z}$. Put $\sigma = |\zeta|^{\frac{1}{2}} \sigma_T$. Then σ is a locally invariant distribution on $\mathfrak{z}_0$ and $\partial(u_{\mathfrak{z}})\sigma = 0$ for $u \in \mathfrak{U}$ (Lemma 17 and Theorem 2). Put $\mathfrak{B} = I(\mathfrak{z}_c)\mathfrak{U}_{\mathfrak{z}}$ where $\mathfrak{U}_{\mathfrak{z}}$ is the space of all elements in $I(\mathfrak{z}_c)$ of the form $u_{\mathfrak{z}}$ $(u \in \mathfrak{U})$. Then it follows from Lemma 19 and condition 2) of Theorem 3, that $\dim I(\mathfrak{z}_c)/\mathfrak{B} < \infty$. Moreover it is clear that $\partial(v)\sigma = 0$ $(v \in \mathfrak{B})$. Since $H_0 \neq 0$ and $\mathfrak{g}$ is semisimple, we know that $\dim \mathfrak{z} < \dim \mathfrak{g}$. Therefore $\sigma = 0$ by induction hypothesis. This proves that $\sigma_T = |\zeta|^{-\frac{1}{2}} \sigma = 0$ and therefore $T = 0$ on Ω_0.

Thus we have now shown that no element of $\mathfrak{g}$, which is not nilpotent, can lie in $\operatorname{Supp} T$. Therefore $\operatorname{Supp} T \subset \mathfrak{N} \cap \Omega$ in the notation of Theorem 4 of [2(e)]. Since the Killing form ω lies in $I(\mathfrak{g}_c)$, we have $\partial(\omega) T = \chi(\omega) T$. Therefore $T = 0$ from Theorem 5 of [2(e)]. This completes the proof of Theorem 3.

THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] N. Bourbaki, *Groupes et Algèbres de Lie, Chapitre I, Algèbres de Lie*, 1960, Hermann, Paris.

[2] Harish-Chandra, (a) "Differential operators on a semisimple Lie algebra," *American Journal of Mathematics*, vol. 79 (1957), pp. 87-120.

(b) "Fourier transforms on a semisimple Lie algebra II," *American Journal of Mathematics*, vol. 79 (1957), pp. 653-686.

(c) "Spherical functions on a semisimple Lie group I," *American Journal of Mathematics*, vol. 80 (1958), pp. 241-310.

(d) "Invariant eigendistributions on semisimple Lie groups," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 117-123.

(e) "Invariant distributions on Lie algebras," *American Journal of Mathematics*, vol. 86 (1964), pp. 271-309.

[3] S. Helgason, *Differential Geometry and Symmetric Spaces*, 1962, Academic Press, New York.

FUNDAMENTAL SOLUTIONS OF INVARIANT DIFFERENTIAL OPERATORS ON SYMMETRIC SPACES.

By S. HELGASON.*

---

1. Introduction and notation. Let D be a differential operator on $\boldsymbol{R}^n$ with constant coefficients. A fundamental solution for D is a distribution T on $\boldsymbol{R}^n$ such that $DT = \delta$ where δ is the delta function on $\boldsymbol{R}^n$. Such a distribution, whose existence is well known, gives by convolution a solution of any equation $Du = f$ on $\boldsymbol{R}^n$, f being any C^∞ function of compact support.

Let D be an arbitrary differential operator on an arbitrary manifold M. In general the differential equation $Du = 0$ has no solution, not even locally. This remains true when M is specialized, unless D is appropriately specialized, too.

In this paper we consider the case when M is a symmetric space of the noncompact type, that is a manifold G/K, where G is a connected semisimple Lie group with center $\{e\}$ and K a maximal compact subgroup. Since a differential operator on $\boldsymbol{R}^n$ with constant coefficients is characterized by its invariance under translations, a natural restriction on the differential operators on G/K is invariance under the action of G. A fundamental solution for such a differential operator D is a distribution T on G/K such that $DT = \delta$, where δ now denotes the unit mass at the point left fixed by K. In this paper the existence of such a fundamental solution is proved (Theorem 4.2; see, however, the introduction to § 4). A fundamental solution for D immediately gives a global solution of the differential equation $Du = f$, where f is a C^∞ function with compact support.

The definition above of a fundamental solution can of course be extended to more general coset spaces. However, for compact spaces, a modification is necessary in order to take into account the nullspace of the operator (cf. § 6).

The proof of Theorem 4.2 leans heavily on Harish-Chandra's theory of spherical functions on the group G. Section 2 summarizes the required results from this theory, combines these with a recent important contribution of Gindikin and Karpelevič, and states some simple consequences. Theorem 3.3 shows that the action of the G-invariant differential operators on G/K on

---

Received December 26, 1963.

* This work was supported by the National Science Foundation, NSF GP-149.

the radial functions on G/K is isomorphic with the action of certain differential operators with constant coefficients. The isomorphism in question is used in Harish-Chandra's work on the Fourier analysis on G ([13], p. 595) and is related to the Radon transform on G/K. For the case when G is complex (§ 5) a more direct isomorphism of this type is given in Theorem 5. 2, again as a consequence of results of Harish-Chandra. As an application we verify the absence of Huygens' principle for the wave equation on G/K, thus supporting the classical conjecture attributed to Hadamard ([6], p. 438), that Laplace's equation on a non-flat Lorentzian manifold does not satisfy Huygens' principle. In § 6 we give a formula for the Green's function (i. e. the fundamental solution of the Laplacian) of a compact, simply connected, semisimple Lie group. In § 7 we investigate whether all the G-invariant differential operators on G/K are induced by the two-sided invariant differential operators on G. It is shown that this is the case when G is a classical group (real or complex) but fails in general for the exceptional groups. However, a weaker result (Theorem 7. 6) is generally true.

The main results of this paper were announced in [16] and [18].

The following notation will be used throughout. As usual $\boldsymbol{R}$ and $\boldsymbol{C}$ denote the fields of real and complex numbers, respectively. The identity of a group will be denoted by e. If A is a group and B a subgroup the set of left cosets aB will be denoted by A/B. For each $x \in A$, the mapping $aB \rightarrow xaB$ of A/B onto itself will be denoted $\tau(x)$; the left translation $a \rightarrow xa$ and the right translation $a \rightarrow ax$ are denoted by $L(x)$ and $R(x)$, respectively. If M is a manifold, $C^\infty(M)$ (respectively $C_c^\infty(M)$) denote the set of C^∞ functions (respectively C^∞ functions of compact support) on M. If $p \in M$, M_p denotes the tangent space to M at p. If τ is a differentiable mapping of M into another manifold N and $p \in M$ then $d\tau_p$ (or simply $d\tau$) denotes the differential of τ at p, mapping M_p into $N_{\tau(p)}$. If τ is a diffeomorphism of M onto N and $f \in C^\infty(M)$ then f^τ denotes the composite function $f \circ \tau^{-1}$. If D is a differential operator on M, $[Df](p)$ (or $D_p(f(p))$) denotes the value of Df at p. If τ is a diffeomorphism of M onto itself then D^τ denotes the differential operator $f \rightarrow (Df^{\tau^{-1}})^\tau$ on M.

If A is a Lie group and B a closed subgroup then $\boldsymbol{D}(A/B)$ denotes the algebra of real linear differential operators on the manifold A/B which are invariant under the action of A. We write $\boldsymbol{D}(A)$ for $\boldsymbol{D}(A/\{e\})$. Lie groups will be denoted by capital Roman letters and their Lie algebras by the corresponding lower case German letters. The adjoint representation of a Lie group A is denoted by Ad_A (or simply Ad). The dual space of a vector space E will usually be denoted $\hat{E}$ or E'.

2. Fourier transforms on symmetric spaces. Let S be a Riemannian globally symmetric space, G the largest connected group of isometries of S in the compact open topology. We assume S is of the noncompact type, that is, G is semisimple and has no compact normal subgroup $\neq \{e\}$. Let o be a fixed point in S, K the (maximal compact) isotropy subgroup of G at o so S can be identified with the left coset space G/K. Let $\mathfrak{k}$ and $\mathfrak{g}$ denote the corresponding Lie algebras, $\mathfrak{p}$ the orthogonal complement of $\mathfrak{k}$ in $\mathfrak{g}$ with respect to the Killing form $\langle\ ,\ \rangle$ of $\mathfrak{g}$. Let $\exp\colon \mathfrak{g} \to G$ denote the exponential mapping for G and $\mathrm{Exp}\colon S_o \to S$ the Exponential mapping for S; the latter is a diffeomorphism. If π denotes the natural mapping of G onto G/K then $d\pi$ maps $\mathfrak{p}$ isomorphically onto S_o and $\pi(\exp X) = \mathrm{Exp}\, d\pi(X)$ for $X \in \mathfrak{p}$. As is customary we identify these two spaces and write $\mathfrak{p}$ for S_o. Let $|X| = \langle X, X\rangle^{\frac{1}{2}}$ for $X \in \mathfrak{p}$. Choose a maximal abelian subspace $\mathfrak{a}$ of $\mathfrak{p}$ (a Cartan subalgebra for S) and put $A = \exp(\mathfrak{a})$. Let $\mathfrak{a}'$ denote the set of regular elements in $\mathfrak{a}$ and put $A' = \exp(\mathfrak{a}')$, $S' = K \cdot \mathrm{Exp}(\mathfrak{a}')$. We fix a Weyl chamber C in $\mathfrak{a}$, i.e. a connected component of $\mathfrak{a}'$. Let $\mathfrak{F}$ denote the set of linear maps of $\mathfrak{a}$ into $\boldsymbol{C}$, so $\mathfrak{F}_{\boldsymbol{R}} = \{\lambda \in \mathfrak{F} \colon \lambda(\mathfrak{a}) \subset \boldsymbol{R}\}$ is the dual space of $\mathfrak{a}$. An element $\lambda \in \mathfrak{F}_{\boldsymbol{R}}$ is called *positive* if $\lambda(H) > 0$ for all $H \in C$. For each $\lambda \in \mathfrak{F}_{\boldsymbol{R}}$ let $\mathfrak{g}_\lambda = \{X \in \mathfrak{g} \colon [H, X] = \lambda(H)X \text{ for } H \in \mathfrak{a}\}$ and let $d_\lambda = \dim(\mathfrak{g}_\lambda)$; λ is called a *root* of S with respect to $\mathfrak{a}$ if $d_\lambda \neq 0$. Let $\Sigma = \{\lambda \in \mathfrak{F}_{\boldsymbol{R}} \colon \lambda > 0, d_\lambda \neq 0\}$, let $2\rho = \Sigma_{\lambda > 0} d_\lambda \lambda$ and let Σ_0 denote the set of elements in Σ which are not integral multiples of other members of Σ. The algebra $\mathfrak{n} = \Sigma_{\lambda > 0}\mathfrak{g}_\lambda$ is a nilpotent subalgebra of $\mathfrak{g}$ and we have Iwasawa decompositions $\mathfrak{g} = \mathfrak{k} + \mathfrak{a} + \mathfrak{n}$, $G = KAN$ where $N = \exp(\mathfrak{n})$. For $g \in G$ let $H(g)$ denote the unique element in $\mathfrak{a}$ satisfying $g \in K \exp H(g) N$. If $h = \exp H$ $(H \in \mathfrak{a})$ we write $H = \log h$.

The Weyl group W of S is the group of linear transformations of $\mathfrak{a}$ induced by those elements of $\mathrm{Ad}_G(K)$ which leave $\mathfrak{a}$ invariant; W also acts on A via the mapping exp and on $\mathfrak{F}_{\boldsymbol{R}}$ by duality.

Let $\boldsymbol{D}_0(G)$ denote the set of elements in $\boldsymbol{D}(G)$ which are invariant under right translations by the elements of K. Then we have a linear mapping μ of $\boldsymbol{D}_0(G)$ onto $\boldsymbol{D}(G/K)$ given by $(\mu(D)f) \circ \pi = D(f \circ \pi)$ for $f \in C^\infty(G/K)$ ([15], p. 390). We have also a homomorphism γ of the algebra $\boldsymbol{D}_0(G)$ onto the algebra $I(\mathfrak{a})$ of W-invariants in the symmetric algebra $S(\mathfrak{a})$ over $\mathfrak{a}$ (Theorem 1, [13], p. 260). The elements in $S(\mathfrak{a})$ will be considered as differential operators on A in the usual manner.

We shall now state the Plancherel formula in the form proved by Harish-Chandra [13] and draw some consequences of a more recent paper of Gindikin-Karpelevič [8].

For $\lambda \in \mathfrak{F}$ let ϕ_λ denote the corresponding spherical function ([13], p. 261) and let $\lambda \to \mathbf{c}(\lambda)$ denote Harish-Chandra's function for S which occurs in the leading term in the asymptotic expansion of ϕ_λ, that is, (cf. [13], p. 283), if dk is the normalized Haar measure on K,

$$\phi_\lambda(\exp H) = \int_K e^{(i\lambda-\rho)(H(\exp H)k)} dk \sim e^{-\rho(H)} \sum_{s \in W} \mathbf{c}(\lambda(sH))\, e^{i\lambda(sH)}$$

for H varying in a subset of the Weyl chamber C and λ varying in a certain subset of $\mathfrak{F}$. The function $\mathbf{c}^{-1}$ is holomorphic in an open subset of $\mathfrak{F}$ containing $\mathfrak{F}_R$ ([13], p. 580).

For a certain space $I_0(G)$, defined below, of complex-valued C^∞ functions on G, bi-invariant under K, the following Plancherel formula holds. For $f \in I_0(G)$, put

$$(1) \qquad \tilde{f}(\lambda) = \int_G f(x)\phi_\lambda(x^{-1})\,dx \qquad (\lambda \in \mathfrak{F}_R).$$

Then ([13], p. 611),

$$(2) \qquad f(x) = w^{-1} \int_{\mathfrak{F}_R} \tilde{f}(\lambda)\phi_\lambda(x) \,|\, \mathbf{c}(\lambda)\,|^{-2}\, d\lambda,$$

$$(3) \qquad \int_G |\, f(x)\,|^2\, dx = w^{-1} \int_{\mathfrak{F}_R} |\, \tilde{f}(\lambda)\,|^2 \,|\, \mathbf{c}(\lambda)\,|^{-2}\, d\lambda,$$

where dx and $d\lambda$ are suitably normalized invariant measures on G and $\mathfrak{F}_R$, respectively, and w is the order of W.

Now let $\alpha \in \Sigma_0$ and consider the subalgebra $\mathfrak{g}^\alpha$ of $\mathfrak{g}$ generated by the subspaces $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$. Let $\mathfrak{k}^\alpha = \mathfrak{g}^\alpha \cap \mathfrak{k}$, $\mathfrak{p}^\alpha = \mathfrak{g}^\alpha \cap \mathfrak{p}$. Then $\mathfrak{g}^\alpha$ is semisimple, $\mathfrak{g}^\alpha = \mathfrak{k}^\alpha + \mathfrak{p}^\alpha$ is a Cartan decomposition and the corresponding symmetric space, say S^α, has rank one. For a Cartan subalgebra of this symmetric space one can choose $\mathfrak{a}^\alpha = \mathfrak{p}^\alpha \cap \mathfrak{a}$. For $\lambda \in \mathfrak{F}$, let λ_α denote its restriction to $\mathfrak{a}^\alpha$. If $\mathbf{c}_\alpha$ denotes the Harish-Chandra function for S^α then, according to Gindikin-Karpelevič [8],

$$(4) \qquad \mathbf{c}(\lambda) = \prod_{\alpha \in \Sigma_0} \mathbf{c}_\alpha(\lambda_\alpha).$$

For each $\alpha \in \Sigma_0$, select $H_\alpha \in \mathfrak{a}$ such that $\langle H_\alpha, H \rangle = \alpha(H)$ for all $H \in \mathfrak{a}$. Then $H_\alpha \in \mathfrak{a}^\alpha$. As shown in [8], $\mathbf{c}(\lambda) = I(i\lambda)/I(\rho)$ where

$$I(\nu) = \prod_{\alpha \in \Sigma} B\left(\frac{p_\alpha}{2}, \frac{p_{\alpha/2}}{4} + \frac{(\nu, \alpha)}{(\alpha, \alpha)}\right),$$

B is the Beta function $(x, y) \to \Gamma(x)\Gamma(y)/\Gamma(x+y)$, and p_α is the multiplicity of α. For the space S^α, Σ_0 consists of just one element, say γ. Then 2γ is the only other possible element in Σ. Denoting their respective multiplicities by p_1 and p_2, the Killing form of $\mathfrak{g}^\alpha$ by $(\ , \ )_\alpha$, the function $\boldsymbol{c}_\alpha(\lambda_\alpha)$ becomes

$$\boldsymbol{c}_\alpha(\lambda_\alpha) = I^{-1}(\rho^\alpha) B\left(\frac{p_1}{2}, i\,\frac{(\lambda_\alpha, \gamma)_\alpha}{(\gamma,\gamma)_\alpha}\right) B\left(\frac{p_2}{2}, \frac{p_1}{4} + i\,\frac{(\lambda_\alpha, 2\gamma)_\alpha}{(2\gamma, 2\gamma)_\alpha}\right),$$

where ρ^α is the function ρ for S^α. If $p_2 = 0$, the last factor does not occur. In any case we have

$$(5) \qquad \boldsymbol{c}_\alpha(\lambda_\alpha) = a_1 \frac{\Gamma(2ia_2\lambda_\alpha(H_\alpha))}{\Gamma(a_3 + 2ia_2\lambda_\alpha(H_\alpha))} \frac{\Gamma(a_4 + ia_2\lambda_\alpha(H_\alpha))}{\Gamma(a_5 + ia_2\lambda_\alpha(H_\alpha))},$$

where the a_i are constants > 0.

Consider now the polynomial function $\pi = \prod_{\alpha \in \Sigma_0} H_\alpha$ on $\mathfrak{F}$ (cf. [13], p. 252). Then

$$(6) \qquad \boldsymbol{c}(\lambda)\pi(\lambda) = \prod_{\alpha\in\Sigma_0} \boldsymbol{c}_\alpha(\lambda_\alpha) H_\alpha(\lambda) = \prod_{\alpha\in\Sigma_0} \boldsymbol{c}_\alpha(\lambda_\alpha)\lambda_\alpha(H_\alpha).$$

Because of (5), $\boldsymbol{c}_\alpha$ has a simple pole at the origin. Thus $\boldsymbol{c}(\lambda)\pi(\lambda)$ is holomorphic and $\neq 0$ for $\lambda = 0$. This means that π_0 in Lemma 29 in [13] p. 580 actually is a constant multiple of π. This shows that the space $I_0(G)$ in [13], pp. 611, 586, consists of all functions ϕ_a of the form

$$(7) \qquad \phi_a(x) = \int_{\mathfrak{F}_R} a(\lambda)\pi(\lambda)\phi_\lambda(x)\,d\lambda,$$

where a belongs to the space $\mathcal{S}(\mathfrak{F}_R)$ of complex-valued rapidly decreasing C^∞ functions on $\mathfrak{F}_R$ ([27], p. 89). Let $\mathcal{I}(\mathfrak{F}_R)$ denote the set of W-invariants in $\mathcal{S}(\mathfrak{F}_R)$. Let $D \in \boldsymbol{D}_0(G)$; then $D\phi_\lambda = \gamma(D)(i\lambda)\phi_\lambda$ ([13], p. 262) so D leaves the space $I_0(G)$ invariant (cf. Theorem 3, [13], p. 586). Let the positive functions ω on $\mathfrak{p}$, Λ and Λ_0 on $\mathfrak{a}$ be determined by

$$(8) \qquad \int_{G/K} f(p)\,dp = \int_{\mathfrak{p}} f(\operatorname{Exp} X)\omega^2(X)\,dX = \int_{\mathfrak{a}} \Lambda(H)\,dH \int_K f(k \exp H \cdot o)\,dk$$

$$(9) \qquad \int_{\mathfrak{p}} F(X)\,dX = \int_{\mathfrak{a}} \Lambda_0(H)\,dH \int_K F(\operatorname{Ad}(k)H)\,dk,$$

f and F being continuous functions of compact support and dp, dX, dH being the volume elements induced by the Killing form. For each integer $q \geqq 0$ let $\omega_q(x) = \omega(X)(1 + |X|)^q$ if $x = k \exp X$ ($k \in K, X \in \mathfrak{p}$). As in [13], p. 585, let $I(G)$ denote the set of all complex-valued C^∞ functions f on G which satisfy the following two conditions: a) f is bi-invariant under K; b) for each $D \in \boldsymbol{D}(G)$ and each integer $q \geqq 0$

$$\tau_{D,q}(f) = \sup_{x\in G} |\omega_q(x)[Df](x)| < \infty.$$

The collection of all seminorms $\tau_{D,q}$ turns $I(G)$ into a locally convex space. Also $I_0(G) \subset I(G)$ and the mapping $a \to \phi_a$ is a continuous mapping of $\mathcal{S}(\mathfrak{F}_R)$ into $I(G)$ (cf. [13], p. 586).

Next we consider the integral transform

$$F_f(h) = e^{\rho(\log h)} \int_N f(hn)\, dn, \qquad f \in I(G), h \in A,$$

dn being a suitably normalized invariant measure on N. This integral transform is used in [13], pp. 595-596 where it is shown that

(i) The mapping $f \to F_f$ is a continuous mapping of $I(G)$ into the space $I(A)$ of W-invariant rapidly decreasing C^∞ functions on A with the usual topology.

(ii)

$$\tilde{f}(\lambda) = \int_A F_f(h)\, e^{-i\lambda(\log h)} dh, \qquad f \in I(G), h \in A.$$

with an appropriate normalization of the invariant measure on A.

The space $I(A)$ is of course an algebra under the convolution product on A.

LEMMA 2.1. *Let $D \in \boldsymbol{D}(G)$ and if $f \in C^\infty(G)$ let*

$$f^\natural(x) = \int_{K\times K} f(kxk')\, dk dk' \qquad (x \in G).$$

Then there exists a constant M_D and finitely many operators $D_1, \cdots, D_{p(D)}$ in $\boldsymbol{D}(G)$ such that the uniform norm $\| \ \|_\infty$ on G satisfies

$$(10) \qquad \| Df^\natural \|_\infty \leqq M_D \sum_{i=1}^{p(D)} \| D_i f \|_\infty \qquad \textit{for all } f \in C_c^\infty(G).$$

Proof. Let $D_1, \cdots, D_p$ be a basis of the subspace of operators in $\boldsymbol{D}(G)$ of order $\leqq$ order (D). If $x \in G$, then $D^{R(x^{-1})} = \sum_{i=1}^{p} a_i(x) D_i$ where $a_i \in C^\infty(G)$. Hence

$$[Df^\natural](x) = \sum_{i=1}^{p} \int_{K\times K} a_i(k') [D_i f](kxk')\, dk dk'$$

and (10) follows.

LEMMA 2.2. *Let $I_c(G) = I(G) \cap C_c^\infty(G)$. Then $I_c(G)$ is a dense subspace of $I(G)$.*

Proof. There exists a sequence $(\phi_i) \subset C_c^\infty(G)$ and an increasing sequence (U_i) of open subsets of G with union G such that

(i) $\phi_i(x) = 1$ for $x \in U_i$;

(ii) For each $D \in \boldsymbol{D}(G)$, $\| D\phi_i \|_\infty \leqq c(D)$ where $c(D)$ is constant which only depends on D.

A proof can be found in Langlands [22], p. 119. Replacing ϕ_i by $\phi_i^\natural$ and U_i by the open set $\{x \in G \colon KxK \subset U_i\}$ we may, by (10), assume that each ϕ_i and each U_i is bi-invariant under K.

Now let $f \in I(G)$, $D \in \boldsymbol{D}(G)$ and q an integer $\geqq 0$. We shall show that $\tau_{D,q}(f - f\phi_i) \to 0$ as $i \to \infty$. It is clear that

$$D(f - f\phi_i) = \sum_{\alpha,\beta} (D_\alpha f) E_\beta (1 - \phi_i) \qquad \text{(finite sum)} \tag{11}$$

where D_α, $E_\beta \in \boldsymbol{D}(G)$. Let $\epsilon > 0$ be given and let N (independent of i) denote the number of terms in (11). Then by (ii) there exists an index i_0 such that for all α, β, i

$$\sup_{x \in G - U_{i_0}} | \omega_q(x) [D_\alpha f](x) [E_\beta(1 - \phi_i)](x) | < \frac{\epsilon}{N}. \tag{12}$$

Since $E_\beta(1 - \phi_i)$ vanishes identically on U_{i_0} if $i \geqq i_0$ it follows from (11) and (12) that

$$\tau_{D,q}(f - f\phi_i) < \epsilon \qquad \text{if } i \geqq i_0.$$

3. The transmutation operator. Since the Weyl group W permutes the elements in $\Sigma \cup (-\Sigma)$ it will also permute the elements in $\Sigma_0 \cup (-\Sigma_0)$. It follows that for each $s \in W$, $\pi^s = \epsilon(s)\pi$ where $\epsilon(s) = \pm 1$. A function a on $\mathfrak{F}_R$ is called *skew* if $a^s = \epsilon(s) a$ for all $s \in W$.

LEMMA 3.1. *Suppose $a \in \mathcal{S}(\mathfrak{F}_R)$ is skew. Then $a = \pi a_1$ where $a_1 \in \mathcal{S}(\mathfrak{F}_R)$.*

Since a is skew it vanishes identically on each hyperplane $H_\alpha(\lambda) = 0$ $(\alpha \in \Sigma)$ in $\mathfrak{F}_R$; (cf. [13], Lemma 10, p. 252). Choose a basis $(x_1, \cdots, x_l)$ of $\mathfrak{a}$ ($=$ dual of $\mathfrak{F}_R$) such that $x_1 = H_\alpha$. Then $a(0, x_2, \cdots, x_l) = 0$ so $a(x_1, \cdots, x_l) = x_1 b(x_1, \cdots, x_l)$ for

$$b(x_1, \cdots, x_l) = \int_0^1 [D_1 a](x_1 s, x_2, \cdots, x_l)\, ds,$$

where D_1 denotes differentiation with respect to the first argument. Thus $a = H_\alpha b$ where $b \in \mathcal{S}(\mathfrak{F}_R)$ and the lemma follows by iteration, since no two elements in Σ_0 are proportional.

PROPOSITION 3.2. *The space $I_0(G)$ is an algebra with respect to the convolution on G.*

Let ϕ_a, ϕ_b be arbitrary elements in $I_0(G)$, where $a, b \in \mathscr{S}(\mathfrak{F}_R)$. Let $A^+ = \exp C$. Then for a suitable normalization of the invariant measure dl on A we have for $z \in G$, ((8), (9) § 2),

$$\phi_a * \phi_b(z) = \int_G \phi_a(zx^{-1})\phi_b(x)\,dx = \int_{A^+} \phi_b(l)\Delta(l) \int_K \phi_a(zkl^{-1})\,dk\,dl.$$

Now

$$(1) \qquad \int_K \phi_a(zkl^{-1})\,dk = \int_K dk \int_{\mathfrak{F}_R} a(\lambda)\pi(\lambda)\phi_\lambda(zkl^{-1})\,d\lambda$$

and since $|\phi_\lambda(x)| \leqq 1$, and $a(\lambda)$ rapidly decreasing we may interchange the order of integration in (1). Using the functional equation for the spherical function it follows that

$$(2) \qquad \phi_a * \phi_b(z) = \int_{A^+} \phi_b(l)\Delta(l) \int_{\mathfrak{F}_R} a(\lambda)\pi(\lambda)\phi_\lambda(z)\phi_\lambda(l^{-1})\,d\lambda\,dl.$$

Let the norm of the linear form $\lambda \in \mathfrak{F}$ be denoted by $|\lambda|$. Now as shown in [13] p. 611 there exists an integer $k \geqq 0$ and, for a fixed b, a positive constant L such that for $\lambda \in \mathfrak{F}_R$

$$|\pi(\lambda)| \int_{A^+} |\phi_b(l)\phi_\lambda(l^{-1})|\,\Delta(l)\,dl \leqq L(1 + |\lambda|)^k.$$

Consequently

$$\int_{A^+}\int_{\mathfrak{F}_R} |\phi_b(l)\Delta(l)a(\lambda)\pi(\lambda)\phi_\lambda(z)\phi_\lambda(l^{-1})|\,d\lambda\,dl$$
$$\leqq \int_{\mathfrak{F}_R} L(1 + |\lambda|)^k\,|a(\lambda)\pi(\lambda)|\,d\lambda$$

which is finite; thus the integrals on the right hand side of (2) can be interchanged so we obtain

$$\phi_a * \phi_b(z) = \int_{\mathfrak{F}_R} a(\lambda)\pi(\lambda)\phi_\lambda(z)\tilde{\phi}_b(\lambda)\,d\lambda.$$

But (i) and (ii) in § 2 show that $\tilde{\phi}_b \in \mathscr{S}(\mathfrak{F}_R)$ so

$$(3) \qquad \phi_a * \phi_b = \phi_c \qquad \text{where } c = a\tilde{\phi}_b$$

and the proposition follows.

THEOREM 3.3. *The mapping $f \rightarrow F_f$ is a topological isomorphism of the algebra $I_0(G)$ onto the algebra $I(A)$. Furthermore,*

$$F_{Df} = \gamma(D)F_f \tag{4}$$

for $D \in \boldsymbol{D}_0(G)$, $f \in I_0(G)$.

LEMMA 3.4. *Let* $f \in I_0(G)$ *and* $D \in \boldsymbol{D}(G)$. *For each compact set* $\Omega \subset G$ *the integral*

$$\int_N [Df](ng)\,dn \tag{5}$$

converges uniformly for $g \in \Omega$.

Writing $g = n_0ak$, where $n_0 \in N$, $a \in A$, $k \in K$, we have with the notation from the proof of Lemma 2.1,

$$\int_N [Df](ng)\,dn = \int_N [(Df)^{R(k^{-1})}](nn_0a)\,dn$$
$$= \sum_{i=1}^{p} a_i(k) \int_N [D_if](nn_0a)\,dn.$$

Therefore it suffices to prove that (5) converges uniformly as g varies in a compact subset of NA. But this is guaranteed by Cor. 2, p. 593 in [13].

In order to prove (4) let $g \in G$ and denote by $A(g)$ the unique element in $\mathfrak{a}$ such that $g \in N \exp A(g)K$. If $h \in A$, the automorphism $n \to hnh^{-1}$ of N has Jacobian at e equal to the determinant of $\mathrm{Ad}(h)$ restricted to $\mathfrak{n}$, i.e., $e^{2\rho(\log h)}$. Hence

$$\int_N f(nh)\,dn = e^{2\rho(\log h)} \int_N f(hn)\,dn \qquad (h \in A) \tag{6}$$

so the function

$$F^*(g) = e^{-\rho(\log A(g))} \int_N f(ng)\,dn$$

is an extension of F_f to a C^∞ function on G. Using Lemma 3.4 we have

$$D_g(e^{\rho(\log A(g))}F^*(g)) = \int_N [Df](ng)\,dn.$$

The function $g \to e^{\rho(\log A(g))}F^*(g)$ is invariant under each substitution $g \to ngk$ $(n \in N, k \in K)$. From Theorem 5.2 in [17] we therefore obtain

$$(D_{\mathfrak{a}})_h(e^{\rho(\log h)}F^*(h)) = \int_N [Df](nh)\,dn \qquad (h \in A)$$

where $D_{\mathfrak{a}} = e^{\rho} \circ \gamma(D) \circ e^{-\rho}$. Since $F_f(h) = F^*(h)$ we obtain (4) by a second application of (6). Next (2) and (ii) in § 2 show immediately that the mapping $f \to F_f$ is one-to-one on $I_0(G)$.

Next we prove that $f \to F_f$ maps $I_0(G)$ onto $I(A)$. Let $\psi \in I(A)$. Then the Fourier transform

$$\hat{\psi}(\lambda) = \int_A \psi(a) e^{-i\lambda(\log a)} da \qquad (\lambda \in \mathfrak{F}_R)$$

is a W-invariant function in $\mathcal{S}(\mathfrak{F}_R)$. Consider now the integral

$$f(x) = w^{-1} \int_{\mathfrak{F}_R} \hat{\psi}(\lambda) \phi_\lambda(x) \,|\, \boldsymbol{c}(\lambda) \,|^{-2} \, d\lambda \qquad (x \in G)$$

whose existence is assured by the following lemma.

LEMMA 3.5. *There exist positive integers a and b such that*

$$\frac{1}{|\,\boldsymbol{c}(\lambda)\,|} \leqq a(1 + |\,\lambda\,|)^b, \qquad \lambda \in \mathfrak{F}_R.$$

In view of (5) and (6) § 2 this is contained in the following lemma.

LEMMA 3.6. *Let a and b be any positive real numbers. The function s on $\boldsymbol{R}$ given by*

$$s(x) = \Gamma(a + ix)(\Gamma(b + ix))^{-1}$$

has each of its derivatives majorized uniformly by a polynomial.

Proof. Using the formula

$$\frac{\Gamma'(z)}{\Gamma(z)} = \sum_1^\infty \frac{1}{n(n+z)} - \gamma - \tfrac{1}{z},$$

where γ is Euler's constant, we find that

$$-i\frac{s'(x)}{s(x)} = -\frac{1}{a+ix} + \frac{1}{b+ix} + (a+ix)\sum_1^\infty \frac{1}{n(n+a+ix)} - (b+ix)\sum_1^\infty \frac{1}{n(n+b+ix)}$$

so

$$s'(x) = s(x)t(x) \tag{7}$$

where $t(x)$ is a function each of whose derivatives is bounded by a polynomial. On the other hand, for each fixed real number c we have for x real,

$$\lim_{|x|\to\infty} |\,\Gamma(c + ix)\,| \exp(\tfrac{1}{2}\pi\,|\,x\,|)\,|\,x\,|^{\frac{1}{2}-c} = (2\pi)^{\frac{1}{2}} \tag{8}$$

(see e.g. [30], p. 151). This equation shows that $s(x)$ is uniformly bounded by a polynomial. Now the lemma follows from (7) by induction.

The formula for $f(x)$ can be written

$$f(x)=w^{-1}\int_{\mathfrak{F}_R}\hat{\psi}(\lambda)\phi_\lambda(x)\,|\,\boldsymbol{c}(\lambda)\,|^{-2}d\lambda=\int_{\mathfrak{F}_R}a(\lambda)\pi(\lambda)\phi_\lambda(x)\,d\lambda,$$

where

$$a(\lambda)=w^{-1}\hat{\psi}(\lambda)\pi(\lambda)\{\boldsymbol{c}(\lambda)\pi(\lambda)\ \operatorname{conj}(\boldsymbol{c}(\lambda)\pi(\lambda))\}^{-1},$$

conj denoting complex conjugation. In view of (5), (6) § 2 and Lemma 3.6 the function a belongs to $\mathcal{S}(\mathfrak{F}_R)$ so

$$f=\phi_a. \tag{9}$$

Now the following lemma implies $\tilde{f}=\hat{\psi}$ and therefore $F_f=\psi$.

Lemma 3.7. *Suppose $a\in\mathcal{S}(\mathfrak{F}_R)$ is skew. Then*

$$\pi(\lambda)\hat{\phi}_a(\lambda)=w\,|\,\boldsymbol{c}(\lambda)\,|^2\,|\,\pi(\lambda)\,|^2a(\lambda).$$

This lemma is an immediate consequence of Theorem 5 in [13] p. 610 (since $\pi_0=\pi$).

Since $(f_1{*}f_2)^{\sim}=\tilde{f}_1\tilde{f}_2$ $(f_1,f_2\in I_0(G))$ formula (ii) § 2 shows that the mapping $f\to F_f$ is a homomorphism of $I_0(G)$ into $I(A)$.

As remarked earlier the mapping $f\to F_f$ is continuous. Thus it just remains to prove that its inverse is continuous. Since $\phi_{s\lambda}=\phi_\lambda$ for $\lambda\in\mathfrak{F}_R$ and $s\in W$ we may, replacing $a(\lambda)$ by $w^{-1}\sum_{s\in W}\epsilon(s)a(\lambda)$ in (7) § 2, always take $a(\lambda)$ skew in the definition of ϕ_a. Since the Euclidean Fourier transform is a homeomorphism between the spaces of rapidly decreasing functions we just have to prove that the mapping $f\to\tilde{f}$ of $I_0(G)$ onto $\mathcal{I}(\mathfrak{F}_R)$ has a continuous inverse. Now $f=\phi_a$ where a is a skew function in $\mathcal{S}(\mathfrak{F}_R)$ which is unique because $\tilde{f}(\lambda)=w\,|\,\boldsymbol{c}(\lambda)\pi(\lambda)\,|^2a(\lambda)\pi(\lambda)^{-1}$. The mapping $h\to|\,\boldsymbol{c}\pi\,|^2h$ is a homeomorphism of $\mathcal{I}(\mathfrak{F}_R)$ onto itself so the question is reduced to the continuity of the mappings

$$\begin{aligned}&d\to\pi d &&\text{of } \mathcal{I}(\mathfrak{F}_R)\text{ into }\mathcal{S}(\mathfrak{F}_R).\\ &a\to\phi_a &&\text{of } \mathcal{S}(\mathfrak{F}_R)\text{ into }I_0(G).\end{aligned}$$

The first is obviously continuous and the continuity of the second was mentioned in § 2. All the statements in Theorem 3.3 have now been verified.

The operator $X\colon f\to F_f$ is a simultaneous "transmutation operator" for the differential operators in $\boldsymbol{D}_0(G)$ in the sense that it transforms a function space on which the $D\in\boldsymbol{D}_0(G)$ act isomorphically onto another function space, where these differential operators act with constant coefficients.

4. Applications to invariant differential operators. At the end of his paper [13], Harish-Chandra states the following two conjectures which would imply the Plancherel formula (3) § 2 for all $f \in I(G)$.

(I) *There exists a polynomial p on* $\mathfrak{F}_R$ *such that* $|c(\lambda)\pi(\lambda)p(\lambda)| \geqq 1$ *for all* $\lambda \in \mathfrak{F}_R$. (Here we have used $\pi = \pi_0$, proved in § 2).

(II) *The mapping* $f \rightarrow F_f$ *is one-to-one on* $I(G)$.

Now (I) is a consequence of formulas (5) and (6) § 2 together with Lemma 3.6. In view of Theorem 3.3, (II) is equivalent to $I_0(G) = I(G)$. On the other hand, the Plancherel formula (3) § 2 for all $f \in I(G)$ together with (ii) § 2 clearly implies (II).

Now I understand that Harish-Chandra has proved the Plancherel formula (3) § 2 for all $f \in I(G)$ in a recent, as yet unpublished, work. In this section we shall therefore assume that $I_0(G) = I(G)$.

If E is any differential operator on G, the adjoint E^* of E (with respect to the measure dx on G) is given by

$$\int_G (Ef)(x)g(x)\,dx = \int_G f(x)(E^*g)(x)\,dx \qquad f, g \in C_c^\infty(G).$$

If T is a distribution on G then ET is by definition the distribution $f \rightarrow T(E^*f)$. If T_1, T_2 are two distributions on G, their convolution $T_1 * T_2$ is the functional on $C_c^\infty(G)$ formally given by (provided the right hand side makes sense)

$$(T_1 * T_2)(f) = \int_{G\times G} f(xy)\,dT_1(x)\,dT_2(y).$$

If at least one T_i has compact support then $T_1 * T_2$ is another distribution on G. For this and several other formal properties of distributions on G, see [9], § 4. In particular if δ_K denotes the distribution $f \rightarrow \int_K f(k)\,dk$ on G we have for $f \in C_c^\infty(G)$

$$f * \delta_K(x) = \int_K f(xk)\,dk, \qquad x \in G.$$

On the space G/K the adjoint operation $D \rightarrow D^*$ on differential operators is defined by means of the G-invariant measure on G/K induced by the invariant measure dx on G and the differentiation of distributions is defined just as above. If T is a distribution on G/K, $\tilde{T}$ denotes the distribution on G given by $\tilde{T}(f) = T(f_\pi)$ where $f \in C_c^\infty(G)$ and f_π is the function on G/K given by $f_\pi \circ \pi = f * \delta_K$. In particular, if $F \in C_c^\infty(G/K)$ then $\tilde{F} = F \circ \pi$ and if δ denotes the distribution $F \rightarrow F(o)$ on G/K then $\tilde{\delta} = \delta_K$. Moreover, the convolution of distributions on G induces an analogous operation $\times$ for distributions on G/K. In fact, if T_1, T_2 are two distributions on G/K, at least one of the compact support, then $T_1 \times T_2$ is the distribution $F \rightarrow (\tilde{T}_1 * \tilde{T}_2)(\tilde{F})$ for $F \in C_c^\infty(G/K)$.

Definition. Let $D \in \boldsymbol{D}(G/K)$. A distribution T on G/K is called a *fundamental solution* for D if $DT = \delta$.

PROPOSITION 4.1. *Let $D \in \boldsymbol{D}(G/K)$ and $F \in C_c^\infty(G/K)$. If T is a fundamental solution for D then the C^∞ function $u = F \times T$ is a solution of the equation*

$$Du = F.$$

Proof. Let $f \in C^\infty(G)$ and let S be a distribution on G. Let f' denote the function $x \rightarrow f(x^{-1})$ and S' the distribution $f \rightarrow S(f')$ $(f \in C_c^\infty(G))$. Let $E \in \boldsymbol{D}(G)$ and let ϵ_E the distribution $f \rightarrow [Ef'](e)$. Then

(1) $Ef = f * \epsilon_E$,

(2) $ES = S * \epsilon_E$.

In fact

$$f * \epsilon_E(x) = \int_G f(xy^{-1})\, d\epsilon_E(y) = \epsilon_E(f^{L(x^{-1})'}) = [Ef^{L(x^{-1})}](e) = [Ef](x).$$

Secondly,

$$\begin{aligned}(ES)(f) = S(E^*f) &= (S * (E^*f)')(e) = (S * (f * \epsilon_{E^*})')(e) \\ &= ((S * (\epsilon_{E^*})') * f')(e) = (S * (\epsilon_{E^*})')(f).\end{aligned}$$

If Y is a left invariant vector field on G then $Y^* = -Y$ so $(\epsilon_{Y^*})' = \epsilon_Y$. If $E_1, E_2 \in \boldsymbol{D}(G)$ then (1) shows that $\epsilon_{E_1E_2} = \epsilon_{E_2} * \epsilon_{E_1}$. Also $(E_1E_2)^* = E_2^*E_1^*$ and $(S_1 * S_2)' = S_2' * S_1'$ if S_1, S_2 are distributions on G with compact support. These formulas combined show that $(\epsilon_{E^*})' = \epsilon_E$ so (2) follows.

Since

$$f * S(x) = \int_G f(xy^{-1})\, dS(y)$$

it follows easily that $(F \times T) \circ \pi = \tilde{F} * \tilde{T}$. Select $E \in \boldsymbol{D}_0(G)$ such that $\mu(E) = D$. Then

$$(D(F \times T)) \circ \pi = E((F \times T) \circ \pi) = (\tilde{F} * \tilde{T}) * \epsilon_E = \tilde{F} * E\tilde{T}$$

Now $\mu(E^*) = (\mu(E))^*$ and $E^* \in \boldsymbol{D}_0(G)$. It follows that if $f \in C_c^\infty(G)$,

$$(E\tilde{T})(f) = \tilde{T}(E^*f) = T((E^*f)_\pi) = T(D^*f_\pi) = DT(f_\pi) = \delta_K(f).$$

Thus

$$(D(F \times T)) \circ \pi = \tilde{F} * \delta_K = F \circ \pi,$$

which proves the proposition.

THEOREM 4.2. *Let D be a nonzero differential operator in $\boldsymbol{D}(G/K)$. Then there exists a fundamental solution for D.*

Consider the transmutation operator $X\colon f \to F_f$ which is a linear homeomorphism of $I(G)$ onto $I(A)$. Its transpose, ${}^tX$, is a linear homeomorphism of the dual space $I(A)'$ onto the dual space $I(G)'$. Now $I(A)'$ can be regarded as the space of W-invariant tempered distributions on the Euclidean space A; in fact if $S \in I(A)'$, the linear mapping $f \to S(w^{-1}\sum_{s\in W} f^s)$ $(f \in C_c^\infty(A))$ is a W-invariant distribution on A, which determines S completely since the space $I_c(A) = I(A) \cap C_c^\infty(A)$ is dense in $I(A)$. This distribution can therefore be denoted by S. Each $D \in S(\mathfrak{a})$, considered as a differential operator on A, with adjoint D^*, acts on $I(A)'$ accordingly: $(DS)(f) = S(D^*f)$. Similarly, the elements of $I(G)'$ can be regarded as distributions on G, bi-invariant under K. In precise terms, let $f \in C_c^\infty(G)$ and let $f^\natural$ be as in Lemma 2.1. If $T \in I(G)'$ we see easily from (10) § 2 that the linear mapping $f \to T(f^\natural)$ is a distribution on G. Since $I_c(G)$ is dense in $I(G)$, (Lemma 2.2), this distribution determines T uniquely and will therefore be denoted by T. Now select $E \in \boldsymbol{D}_0(G)$ such that $\mu(E) = D$. The distribution δ_K clearly belongs to $I(G)'$ so there exists a unique element $\epsilon \in I(A)'$ such that ${}^tX\epsilon = \delta_K$. By theorem of Hörmander [19] and Łojasiewicz [23] each nonzero differential operator with constant coefficients on $\boldsymbol{R}^n$ maps the space of tempered distributions onto itself. In particular there exists a tempered distribution S on A such that

$$\gamma(E^*) * S = \epsilon.$$

Both $\gamma(E^*)^*$ and ϵ are W-invariant so we can also take S invariant under W, that is, $S \in I(A)'$. Put $T_1 = {}^tX(S)$; then $T_1 \in I(G)'$ and if $f \in C_c^\infty(G)$ we have

$$T_1(E^*f) = T_1((E^*f)^\natural) = {}^tXS(E^*f^\natural) = S(XE^*f^\natural) = S(\gamma(E^*)Xf^\natural)$$
$$= (\gamma(E^*)^*S)(Xf^\natural) = ({}^tX\epsilon)(f^\natural) = \delta_K(f^\natural) = \delta_K(f).$$

This proves that

$$ET_1 = \delta_K. \tag{3}$$

The proof of Theorem 4.2 is now easily finished. Let T denote the linear functional

$$F \to T_1(F \circ \pi) \qquad F \in C^\infty(G/K).$$

Then T is a distribution on G/K and $T_1 = \tilde{T}$. Also

$$\begin{aligned}(DT)(F) &= T(D^*F) = T_1((D^*F)\circ\pi) \\ &= T_1(E^*(F\circ\pi)) = (ET_1)(F\circ\pi) = \delta_K(F\circ\pi) = F(o).\end{aligned}$$

Thus

$$DT = \delta$$

and the theorem is proved. The proof of (3) has the following corollary.

COROLLARY 4.3. *Let E be any nonzero element in $\boldsymbol{D}_0(G)$. Then*

(i) *E maps $I(G)'$ onto itself.*

(ii) *E is a one-to-one mapping of $I(G)$ into itself.*

The first part is proved exactly as formula (3). For the second part, suppose $f \in I(G)$ such that $Ef = 0$. Then $X(Ef) = \gamma(E)Xf = 0$. Since Xf and $\gamma(E)Xf$ lie in $I(A)$ we obtain by taking Fourier transform, $Xf = 0$ so $f = 0$.

PROPOSITION 4.4. *Each $D \neq 0$ in $\boldsymbol{D}(G/K)$ is a one-to-one mapping of $C_c^\infty(G/K)$ into itself.*

Proof. Suppose $F \in C_c^\infty(G/K)$ satisfies $DF = 0$. Let

$$f^*(g,x) = \int_K F(gkx\cdot o)\,dk, \qquad g, x \in G,$$

and select $E \in \boldsymbol{D}_0(G)$ such that $\mu(E) = D$. Then by the Darboux equation [15], p. 443

$$E_x(f^*(g,x)) = \int_K [DF](gkx\cdot o)\,dk = 0. \tag{4}$$

For each fixed g, the function $x \to f^*(g,x)$ is an element of $I(G)$. By Cor. 4.3 (ii) and (4), $f^*(g,x) = 0$ for all $x \in G$. Putting $x = e$, it follows that $F = 0$.

5. The case of a complex group G.

In this section we shall assume that S is a Riemannian globally symmetric space of the noncompact type for which the group G is a complex Lie group. In this case we give an inversion formula for the transmutation operator $X: f \to F_f$ and a representation formula for the differential operators in $\boldsymbol{D}(G/K)$, which leads to a sharper version of Theorem 4.2 and certain applications.

THEOREM 5.1. *Let $\tilde{\square}$ be any element in $\boldsymbol{D}_0(G)$ such that $\gamma(\tilde{\square}) = \pi^2$. Then the inverse of the operator $f \to F_f$ of $I(G)$ onto $I(A)$ is given by*

$$(1)\qquad \bar{\square}_g\left(\int_K F_f(\exp(H(gk)))\, e^{-\rho(H(gk))}\,dk\right) = cf(g)$$

where c is a constant $\neq 0$, independent of f.

Proof. Suppose first that f has compact support. Then

$$\int_K F_f(\exp(H(gk)))\, e^{-\rho(H(gk))}\,dk = \int_K dk \int_N f(\exp(H(gk))n)\,dn$$
$$= \int_K \int_N f(gkn)\,dn\,dk,$$

since dn is invariant and f is invariant under left translations from K. The last integral equals $I_f(g)$ in the notation of [17], § 6 and (1) follows from Theorem 6.1 in [17]. Now if $f \in I(G)$ is arbitrary there exists a sequence (f_n) of elements in $I_c(G)$ which converges to f (in $I(G)$). Then the inequality in Cor. 2 of [13], p. 593, shows that for each $D \in \boldsymbol{D}(A)$, DF_{f_n} converges to DF_f uniformly on each compact subset of A. For $a \in A$ the left hand side of (1) can be written

$$\delta'(\bar{\square})_a \int_K F_f(\exp(H(ak)))\, e^{-\rho(H(ak))}\,dk$$

where $\delta'(\bar{\square})$ is a certain differential operator on A' (cf. [13], Lemma 23, p. 265). From these statements it follows that (1) is valid on A' for each $f \in I(G)$. But $KA'K$ is dense in G so (1) holds on the entire G by continuity.

A function F on $S = G/K$ (or $\mathfrak{p}$) will be called a *radial function* if $F^{\tau(k)} = F$ (or $F^{\mathrm{Ad}(k)} = F$) for all $k \in K$. Similarly radial distributions are defined. Let $I(\mathfrak{p})$ denote the algebra of radial polynomial functions on $\mathfrak{p}$. If we combine the isomorphism of $\boldsymbol{D}(G/K)$ onto $I(\mathfrak{a})$ ([13], Theorem 1, [15], p. 432) with Chevalley's restriction isomorphism ([15], p. 430) of $I(\mathfrak{p})$ onto $I(\mathfrak{a})$ we obtain an isomorphism λ of $\boldsymbol{D}(G/K)$ onto $I(\mathfrak{p})$.

THEOREM 5.2. *Let F be any radial C^∞ function on G/K and $D \in \boldsymbol{D}(G/K)$. Then*

$$(2)\qquad (Df)\circ \mathrm{Exp} = \frac{1}{\omega}\,\lambda(D)\,(\omega(F\circ \mathrm{Exp})).$$

Proof. Let $\bar{F}$ denote the function on $\mathfrak{a}$ given by $\bar{F}(H) = F(\mathrm{Exp}\,H)$. Then Lemma 55, p. 305 in [13] (cf. also [3], Theorem 28), can be stated

$$(3)\qquad (DF)^- = \Delta^{-\frac{1}{2}}\Gamma(D)\,(\Delta^{\frac{1}{2}}\bar{F}) \qquad (\text{on } \mathfrak{a}').$$

and Theorem 1 in [12] implies

$$(4) \qquad (\lambda(D)(F \circ \operatorname{Exp}))^{-} = \Lambda_0^{-\frac{1}{2}}\Gamma(D)(\Lambda_0^{\frac{1}{2}}\bar{F}) \qquad (\text{on } \mathfrak{a}').$$

Now $\bar{\omega}^2 = \Lambda/\Lambda_0$ so by (3) and (4)

$$(DF)^{-} = \Lambda^{-\frac{1}{2}}\Gamma(D)(\Lambda^{\frac{1}{2}}\bar{F}) = \Lambda^{-\frac{1}{2}}\Gamma(D)(\Lambda_0^{\frac{1}{2}}\bar{\omega}\bar{F}) = \bar{\omega}^{-1}(\lambda(D)(\omega(F \circ \operatorname{Exp}))^{-}.$$

This proves formula (2) on the set $\mathfrak{a}'$. By continuity, it holds on $\mathfrak{a}$. But both sides of (2) are radial functions so (2) holds on the entire space $\mathfrak{p}$.

Remark. I am indebted to B. Kostant for conversations which raised the question of a formula of the type of (2). For the bi-invariant differential operators on any real semisimple Lie group a local formula similar to (2) has also been proved by Harish-Chandra.

As a consequence of Theorem 5.2 we can now prove a sharper form of Theorem 4.2.

THEOREM 5.3. *Let $D \in \mathbf{D}(G/K)$. The radial fundamental solutions Q for D are the distributions*

$$Q\colon F \rightarrow Q_0(\omega(F \circ \operatorname{Exp})), \qquad F \in C_c^{\infty}(G/K),$$

where Q_0 is an arbitrary radial fundamental solution for $\lambda(D^*)^*$.

Proof. Let $I_c^{\infty}(G/K)$ and $I_c^{\infty}(\mathfrak{p})$ denote the spaces of radial C^{∞} functions on G/K and $\mathfrak{p}$, respectively, equipped with the topologies induced by $C_c^{\infty}(G/K)$ and $C_c^{\infty}(\mathfrak{p})$. Their duals can be indentified (compare § 4) with the spaces of radial distributions on G/K and $\mathfrak{p}$, respectively. The mapping

$$Y\colon F \rightarrow \omega(F \circ \operatorname{Exp})$$

is a linear homeomorphism of $I_c^{\infty}(G/K)$ onto $I_c^{\infty}(\mathfrak{p})$ and according to (2),

$$(5) \qquad YDF = \lambda(D)YF$$

so Y is a simultaneous transmutation operator for the differential operators in $\mathbf{D}(G/K)$. The transpose ${}^tY$ is a linear homeomorphism between the mentioned distribution spaces. In particular, the distribution $({}^tY)^{-1}(\delta)$ is the delta distribution $\delta_0\colon f \rightarrow f(0)$ on $\mathfrak{p}$. Now let Q_0 be any radial distribution on $\mathfrak{p}$. Then the equation

$$\lambda(D^*)^*Q_0 = \delta_0$$

is easily seen (by (5)) to be equivalent to the equation

$$D({}^tYQ_0) = \delta;$$

this proves the theorem.

Examples.

I. The Laplace-Beltrami operator on S. Let Δ and Δ_0 denote the Laplace-Beltrami operators on S and $\mathfrak{p}$ respectively. As pointed out in [15], p. 451, Δ is the image under the mapping μ of the Casimir operator in $\boldsymbol{D}(G)$. Therefore, if we use Cor. 2, p. 271 in [13], we see that $\lambda(\Delta) = \Delta_0 - |\rho|^2$. Now $\Delta^* = \Delta$, $\Delta_0{}^* = \Delta_0$, and a well known fundamental solution of $\Delta_0 - |\rho|^2$ is given by

$$(6) \qquad Q_0(X) = 2^{-\frac{1}{2}(n+2)}\pi^{\frac{1}{2}(1-n)}(i\,|\rho|\,|X|^{-1})^{\frac{1}{2}(n-2)}N_{\frac{1}{2}(n-2)}(i\,|\rho|\,|X|),$$

where $n = \dim(\mathfrak{p})$ and N_p is the Neumann function (see e. g. Schwartz [28], Exposé 8). Thus $|X|^{n-2}Q_0(X)$ is bounded near the origin $X = 0$ so Q_0 is a locally integrable function on $\mathfrak{p}$.

THEOREM 5.4. *The Laplace-Beltrami operator Δ on S has a locally integrable fundamental solution Q given by*

$$(7) \qquad Q \circ \operatorname{Exp} = \omega^{-1}Q_0,$$

where Q_0 is given by (6).

In fact, the fundamental solution Q for Δ given by Theorem 5.2 is locally integrable on S and satisfies for all $F \in C_c^\infty(S)$

$$Q(F) = \int_S Q(p)F(p)\,dp = \int_{\mathfrak{p}} Q(\operatorname{Exp} X)F(\operatorname{Exp} X)\omega^2(X)\,dX,$$

$$Q(F) = Q_0(\omega(F \circ \operatorname{Exp})) = \int_{\mathfrak{p}} Q_0(X)\omega(X)\,(F \circ \operatorname{Exp})(X)\,dX,$$

so $Q \circ \operatorname{Exp} = \omega^{-1}Q_0$ as stated.

Remark. For the case when S is the space of positive definite Hermitian $m \times m$ matrices a formula of the type (7) was given without proof by Dynkin [7].

II. The wave equation on S. Absence of Huygens' Principle. In [17] the solution of the Cauchy problem for the wave equation $(\Delta - \partial^2/\partial t^2)u = 0$ on S is written down. However, the solution formula, although explicit, seems too complicated to settle immediately the question whether Huygens' principle[1] holds for the wave equation on S. This can be answered in the negative by using the representation of Δ in Example I.

---

[1] For a discussion and results concerning Hadamard's problem about the occurrence of Huygen's principle, see e. g. Hadamard [10], p. 236, Stellmacher [26], Asgeirsson [2], Helgason [14], p. 296.

According to Hadamard, Huygens' principle cannot hold if $n = \dim S$ is even. Hence we may assume n odd. The equation

$$(\Delta_0 - |\rho|^2 - \partial^2/\partial t^2)u = 0$$

for damped waves in the Euclidean space $\mathfrak{p}$ has in the interior of the retrograde light cone

$$D_0 = \{(X,t) \in \mathfrak{p} \times \boldsymbol{R} : t^2 - |X|^2 > 0, t < 0\}$$

a solution of the form

$$u(X,t) = U(X,t)(t^2 - |X|^2)^{\frac{1}{2}(1-n)} + V(X,t)\log(t^2 - |X|^2), \tag{8}$$

where the functions U and V are bounded for small X and t and $V(0,0) \neq 0$ (cf. Hadamard [10], p. 107). Replacing X in (8) by $\mathrm{Ad}(k)X$ and integrating over K we may assume that U and V are radial functions in the variable X. Now consider the function F given by $F(\mathrm{Exp}\,X,t) = \omega(X)^{-1}u(X,t)$ for $(X,t) \in D_0$. Since $\lambda(\Delta) = \Delta_0 - |\rho|^2$ we have from (2)

$$\begin{aligned}[(\Delta - \frac{\partial^2}{\partial t^2})F](\mathrm{Exp}\,X,t) &= [\Delta F](\mathrm{Exp}\,X,t) - \frac{\partial^2}{\partial t^2}F(\mathrm{Exp}\,X,t) \\ &= \omega(X)^{-1}[(\Delta_0 - |\rho|^2)u](X,t) - \omega(X)^{-1}\frac{\partial^2}{\partial t^2}u(X,t) = 0\end{aligned}$$

so the function F satisfies the wave equation in the interior of the retrograde light cone

$$\boldsymbol{D}_0 = \{(\mathrm{Exp}\,X,t) \in S \times \boldsymbol{R} : (X,t) \in D_0\}.$$

Let r denote the Lorentzian distance of $(\mathrm{Exp}\,X,t)$ from the origin in $S \times \boldsymbol{R}$, i.e., $r(\mathrm{Exp}\,X,t) = (t^2 - |X|^2)^{\frac{1}{2}}$. Then F is a solution of the wave equation in $\boldsymbol{D}_0$ of the form

$$F(p) = U_1(p)r(p)^{1-n} + V_1(p)\log r(p) \qquad (p \in \boldsymbol{D}_0)$$

where $V_1(0) \neq 0$. This means that Huygens' principle is absent (cf. Hadamard [10], p. 236, Courant-Hilbert [6], p. 438).

6. The Green's function of a compact semisimple Lie group. Let M be a compact Riemannian manifold of dimension m, $C_0^\infty(M)$ the subspace of $f \in C^\infty(M)$ which satisfy $\int_M f(x)\,dx = 0$, dx denoting the Riemannian measure on M. The Laplace-Beltrami operator Δ on M is an isomorphism of $C_0^\infty(M)$ onto itself. The inverse operator, say Γ, is an integral operator whose kernel, *the Green's kernel of M*, is a differentiable function outside the diagonal of $M \times M$ and is characterized by the following conditions (see de Rham [24], p. 157):

(i) $g(x,y) = g(y,x)$ for $x \neq y$ in M.

(ii) $d(x,y)^{m-2} g(x,y)$ is bounded, d denoting the distance on M. (If $m=2$ this condition should read "$(\log d(x,y))^{-1} g(x,y)$ is bounded").

(iii) $[\Gamma f](x) = \int_M g(x,y) f(y)\, dy$ for all $f \in C_0^\infty(M)$.

(iv) $\int_M g(x,y)\, dy \equiv 0$.

Now suppose M is globally symmetric and let U be the largest connected group of isometries of M in the compact open topology. Then U is compact. Fix a point $o \in M$ and let K be the isotropy subgroup of U at o, so $M = U/K$. If g is the Green's kernel of M it is easy to see that for each $u \in U$, the function $(x,y) \to g(u\cdot x, u\cdot y)$ also satisfies (i)-(iv) so $g(x,y) \equiv g(u\cdot x, u\cdot y)$. Now define the function γ on $M - \{o\}$ by $\gamma(x) = g(x,o)$. Then $\gamma(k\cdot x) = \gamma(x)$ for $k \in K$, $x \neq o$ in M. The function γ on $M - \{o\}$ will be called *the Green's function of* M. Because of (ii), γ is an integrable function on M. Let $|M|$ denote the total volume of M and δ the distribution $f \to f(o)$ on M.

THEOREM 6.1. *Let M be a compact Riemannian globally symmetric space. The Green's function γ, considered as a distribution on M, is characterized as the unique solution of the equations*

$$\Delta\gamma = \delta - \frac{1}{|M|}, \qquad \gamma(1) = 0. \tag{1}$$

For each function f on $M = U/K$ let $\tilde{f} = f \circ \pi$, where $\pi\colon U \to U/K$ is the natural mapping. We normalize the invariant measure du on U such that $|U| = \int_U du$ equals $|M|$. For two distributions T_1, T_2 on M, $T_1 \times T_2$ is defined as in § 4. Then if $f \in C^\infty(M)$ and $u \in U$,

$$\begin{aligned}(f \times \gamma)(u\cdot o) &= \int_U \tilde{f}(v)\tilde{\gamma}(v^{-1}u)\, dv = \int_U \tilde{f}(v) g(v^{-1}u\cdot o, o)\, dv \\ &= \int_U f(v\cdot o) g(u\cdot o, v\cdot o)\, dv = \int_M f(x) g(u\cdot o, x)\, dx \\ &= \int_M g(u\cdot o, x) f_0(x)\, dx = [\Gamma f_0](u\cdot o),\end{aligned}$$

where we have put $f_0 = f - |M|^{-1}\int f(x)\, dx$. It follows that

$$\Delta(f \times \gamma) = f - \frac{1}{|M|}\int f(x)\,dx \qquad (f \in C^\infty(M)). \tag{2}$$

Now there exists an operator $\tilde{\Delta} \in \boldsymbol{D}(U)$, invariant under all right translations by elements of K, such that $\tilde{\Delta}(f \circ \pi) = \Delta f \circ \pi$ for all $f \in C^\infty(M)$ (cf. [15], p. 390). Computations, analogous to those of the proof of Prop. 4.1 show that ($*$ denoting convolution on U),

$$(\Delta(f \times \gamma)) \circ \pi = \tilde{f} * \tilde{\Delta}\tilde{\gamma}, \qquad f \in C^\infty(M)$$
$$(\Delta\gamma)^{\sim} = \tilde{\Delta}\tilde{\gamma}$$

so

$$\Delta(f \times \gamma) = f \times \Delta\gamma \qquad f \in C^\infty(M). \tag{3}$$

Next let s denote the geodesic symmetry of M with respect to o and let σ denote the automorphism $u \rightarrow sus$ of U. Let $f^\natural$ denote the function on M given by the average,

$$f^\natural(x) = \int_K f(k \cdot x)\,dk$$

and if F is a function on U let F' denote the function $u \rightarrow F(u^{-1})$ on U. If F is bi-invariant under K then $F' = F^\sigma$ since $\sigma(u) \in Ku^{-1}K$ for $u \in U$. Since $s \circ \pi = \pi \circ \sigma$ (cf. [15], p. 174) it is easy to verify $(((f^s)^\natural)^{\sim})^\sigma = (f^\natural)^{\sim}$ for $f \in C^\infty(M)$. Consequently, by (2) and (3),

$$\begin{aligned}(f \times \Delta\gamma)(o) &= (f^s \times \Delta\gamma)(o) = ((f^s)^\natural \times \Delta\gamma)(o)\\ &= (((f^s)^\natural)^{\sim} * (\Delta\gamma)^{\sim})(e) = (\Delta\gamma)^{\sim}((((f^s)^\natural)^{\sim})') = (\Delta\gamma)^{\sim}((f^\natural)^{\sim})\\ &= \Delta\gamma(f^\natural) = \Delta\gamma(f)\end{aligned}$$

so

$$(f \times \Delta\gamma)(o) = \Delta\gamma(f) \qquad f \in C^\infty(M). \tag{4}$$

Now (1) results from formulas (2)-(4). On the other hand, suppose γ_1 is another distribution satisfying (1). Then $\gamma - \gamma_1$ satisfies the elliptic equation $\Delta(\gamma - \gamma_1) = 0$. By the regularity theorem for elliptic operators (see e.g. [19], p. 178) $\gamma - \gamma_1$ is a C^∞ function on M, hence a constant which must be 0, because $\gamma(1) = \gamma_1(1) = 0$.

In the case when M has rank one, γ is determined in more explicit terms in [15], p. 447.

Suppose now U is a compact, simply connected semisimple Lie group with Riemannian structure induced by the negative of the Killing form $\langle\ ,\ \rangle$ of the Lie algebra $\mathfrak{u}$ of U. We shall compute the Green's function of U. Let $\mathfrak{t}$ be any maximal abelian subalgebra of $\mathfrak{u}$ and let $T = \exp(\mathfrak{t})$; T is taken with the Riemannian structure induced by that of U. Let $\mathfrak{t}_e$ denote the unit lattice $\{H \in \mathfrak{t}: \exp H = e\}$ in $\mathfrak{t}$ and let Λ_e denote the lattice of complex-valued linear functions λ on $\mathfrak{t}$ satisfying $\lambda(\mathfrak{t}_e) \subset 2\pi i \boldsymbol{Z}$, where $\boldsymbol{Z}$ denotes the set of integers. For each $\lambda \in \Lambda_e$, the function e^λ on T determined by $e^\lambda(\exp H) = e^{\lambda(H)}$ $(H \in \mathfrak{t})$ is a character of T. On the other hand, each character χ of T arises in this manner from a unique $\lambda \in \Lambda_e$ so Λ_e can be identified with the character group of T. If f is a function on U, let $\bar{f}$ denote its restriction to T.

The superscript denoting complexification we consider the set of roots of the complex Lie algebra $\mathfrak{u}^c$ with respect to the Cartan subalgebra $\mathfrak{t}^c$. Suppose the real vector space spanned by the roots is ordered by means of some Weyl chamber C in $i\mathfrak{t}$ (cf. § 2) and let ρ denote one half of the sum of the positive roots. Let W denote the Weyl group of U; it acts on T and on $\mathfrak{t}$. If $s \in W$ let $\epsilon(s)$ denote the determinant of s acting on $\mathfrak{t}$. It is well known that $\rho \in \Lambda_e$ and $s(\mathfrak{t}_e) \subset \mathfrak{t}_e$ for each $s \in W$ so the function $j = \sum_{s \in W} \epsilon(s) e^{s\rho}$ is a well-defined function on T.

The extension of $\langle\ ,\ \rangle$ to a bilinear form on $\mathfrak{u}^c \times \mathfrak{u}^c$ is the Killing form of $\mathfrak{u}^c$. If λ is a linear form on $\mathfrak{t}^c$, let $H_\lambda \in \mathfrak{t}^c$ be determined by $\langle H_\lambda, H\rangle = \lambda(H)$ for $H \in \mathfrak{t}^c$. If $H_\lambda \in i\mathfrak{t}$ we put $|\lambda| = \langle H_\lambda, H_\lambda\rangle^{\frac{1}{2}}$. Then $H_\rho \in C$ since $\rho > s\rho$ for $s \neq e$ in W.

LEMMA 6.2. *The orbit of* ρ *under* W *coincides with the set of elements in* Λ_e *of norm* $|\rho|$, i.e.,

$$\{s\rho : s \in W\} = \{\lambda \in \Lambda_e : |\lambda| = |\rho|\}.$$

Proof. Consider the subring A of $C^\infty(T)$ generated by the functions e^μ $(\mu \in \Lambda_e)$. A function $f \in C^\infty(T)$ is called *skew* if $f^s = \epsilon(s) f$. It is well known and easily proved (cf. Lemma 3.1) that if $f \in C^\infty(T)$ is skew then $f = j f_1$ where $f_1 \in C^\infty(T)$ and if $f \in A$ then $f_1 \in A$. Now suppose $\lambda \in \Lambda_e$ and $|\lambda| = |\rho|$. Then

$$\sum_{s \in W} \epsilon(s) e^{s\lambda} = \sum_{s \in W} \epsilon(s)^{s\rho} \sum_{\mu} a_\mu e^\mu.$$

Now suppose $a_\mu \neq 0$ for some $\mu > 0$ in Λ_e. Then if λ_0 is the highest among the linear forms $s\lambda$ $(s \in W)$ we have $\lambda_0 = \rho + \mu$. But $\mu > 0$ implies $\langle \mu, \rho\rangle = \mu(H_\rho) > 0$ so $|\lambda_0|^2 > |\rho|^2 + |\mu|^2$ which is a contradiction. Similarly, $a_\mu = 0$ if $\mu < 0$ so we find that $\lambda_0 = \rho$, proving the lemma.

To any distribution S on T we associate the Fourier series

$$S \sim \sum_{\lambda \in \Lambda_e} a(\lambda) e^\lambda, \qquad a(\lambda) = \frac{1}{|T|} S(e^{-\lambda})$$

in the usual manner ([27]). Since j divides each skew function in $C^\infty(T)$ it is easily seen that the linear form J on $C^\infty(T)$ given by

$$J(F) = \frac{1}{|U|}(-1)^{\frac{1}{2}(\dim U - \dim T)}\Big(\frac{1}{j}\sum_{s \in W} \epsilon(s) F^s\Big)(e)$$

for $F \in C^\infty(T)$ is a distribution on T. The distribution γ is an integrable function on U and being invariant under all automorphisms of U is completely determined by its restriction $\bar{\gamma}$ to T. Let T' denote the set of points $t \in T$ for which $j(t) \neq 0$.

THEOREM 6.3. *The Green's function γ on U is given by*

$$\bar{\gamma} = \frac{1}{j}(\gamma_\rho * J) \qquad \text{on } T',$$

where γ_ρ is the distribution on T given by the Fourier series

$$\gamma_\rho = \sum_{|\lambda| \neq |\rho|} \frac{1}{|\lambda|^2 - |\rho|^2} e^\lambda.$$

Proof. Let $I(U)$ denote the space of all $f \in C^\infty(U)$ which are invariant under all automorphisms of U; the space $C^\infty(U)$ is given the usual Schwartz topology and $I(U)$ is a closed subspace. Let $S(T)$ denote the space of skew functions in $C^\infty(T)$, topologized similarly. The mapping $X: f \to j\bar{f}$ is a linear homeomorphism of $I(U)$ onto $S(T)$ and

$$X(\Delta f) = (\Delta_T + |\rho|^2) X f \qquad f \in I(U),$$

where Δ is the Laplacian on U and Δ_T the Laplacian on T. This formula can be obtained from Weyl's classical character formula and also from Berezin [3], Theorem 2.4 or Harish-Chandra [13], p. 271. The transpose ${}^tX$ is a linear one-to-one mapping of the dual $S(T)'$ (which can be identified with the space of skew distributions on T) onto the dual $I(U)'$ (which can be regarded as the space of invariant distributions on U). If we put

$$Y_0 = ({}^tX)^{-1}\Big(\delta - \frac{1}{|U|}\Big)$$

and find a solution $S_0 \in S(T)'$ to the equation

$$(\Delta_T + |\rho|^2) S_0 = Y_0$$

then $\gamma = {}^tX(S_0)$. Using Weyl's integral formula

$$\frac{1}{|U|}\int_U f(u)\,du = \frac{1}{w}\frac{1}{|T|}\int_T |j(t)|^2 \bar{f}(t)\,dt \qquad (w = \text{order of } W) \tag{5}$$

Y_0 is easily found whereupon S_0 can be written down as a Fourier series on T. Using (5) once more, and Lemma 6.2, the theorem follows. We omit the routine details of the computation.

7. Restriction of the central operators in $\boldsymbol{D}(G)$.

Returning now to the notation of § 1 and § 2 let $\boldsymbol{Z}(G)$ denote the center of $\boldsymbol{D}(G)$. As is well known, $\boldsymbol{Z}(G)$ consists of all differential operators on G which are invariant under all right and all left translations on G. We shall now consider the question whether $\boldsymbol{Z}(G)$ is mapped onto $\boldsymbol{D}(G/K)$ by means of the natural homomorphism μ of $\boldsymbol{D}_0(G)$ onto $\boldsymbol{D}(G/K)$ given by $(\mu(D)f)\circ\pi = D(f\circ\pi)$ for all $D \in \boldsymbol{D}_0(G)$, $f \in C^\infty(G/K)$.

At the same time we shall consider another homomorphism of $\boldsymbol{Z}(G)$ into $\boldsymbol{D}(G/K)$. Since G is a Lie transformation group of G/K we have a representation ν of G on $C^\infty(G/K)$ by putting $\nu(g)f = f^{\tau(g)}$ for $g \in G$, $f \in C^\infty(G/K)$ and $\tau(g)$ as in § 1. This induces a representation of $\mathfrak{g}$ on $C^\infty(G/K)$ which extends to a homomorphism, also denoted ν, of $\boldsymbol{D}(G)$ into the algebra of all differential operators on G/K.

Lemma 7.1. *Let $D \in \boldsymbol{D}(G)$, $g \in G$. Then*

$$\nu(D)\nu(g) = \nu(g)\nu(D^{R(g)}).$$

Proof. It suffices to prove this for $D = X \in \mathfrak{g}$. If $f \in C^\infty(G)$ we have for $p \in G/K$,

$$\begin{aligned}[\nu(X)\nu(g)f](p) = [\frac{d}{dt}f(g^{-1}\exp(-tX)\cdot p)](0) &= [\nu(\mathrm{Ad}(g^{-1})X)f](g^{-1}\cdot p)\\ &= [\nu(g)\nu(X^{R(g)})f](p).\end{aligned}$$

Lemma 7.2. *Let $D \in \boldsymbol{D}(G)$, $f \in C^\infty(G)$. Then*

$$[\nu(D)f](\pi(g)) = [(D^*)^{R(g)}(f\circ\pi)](g),$$

where D^* is the adjoint of D.

Proof. If $g \in G$ and $X \in \mathfrak{g}$ we have

$$\begin{aligned}[\nu(X)f](\pi(g)) &= \{\frac{d}{dt}(f(\exp(-tX)\cdot\pi(g)))\}_{t=0}\\ &= \{\frac{d}{dt}((f\circ\pi)(g\exp(\mathrm{Ad}(g^{-1})(-tX))))\}_{t=0}\end{aligned}$$

so

$$[\nu(X)f](\pi(g)) = [\mathrm{Ad}(g^{-1})(-X)(f\circ\pi)](g). \tag{1}$$

Let $X_1, \cdots, X_r \in \mathfrak{g}$. We shall then prove by induction on r that

$$(2) \qquad [\nu(X_1 \cdots X_r)f](\pi(g)) = (-1)^r [\mathrm{Ad}(g^{-1})[X_r \cdots X_1)(f \circ \pi)](g).$$

For $r = 1$ this is proved and assuming (2) for a fixed r we have for $X_0 \in \mathfrak{g}$

$$\begin{aligned}
[\nu(X_0 \cdots X_r)f](\pi(g)) &= [\nu(X_0)(\nu(X_1 \cdots X_r)f)](\pi(g)) \\
&= \{\frac{d}{dt}(\nu(X_1 \cdots X_r)f(\pi(\exp(-tX_0)g)))\}_{t=0} \\
&= (-1)^r \{\frac{d}{dt}([(\mathrm{Ad}(g^{-1})\mathrm{Ad}(\exp(tX_0)(X_r \cdots X_1))(f \circ \pi)] \\
&\qquad\qquad (g \exp \mathrm{Ad}(g^{-1})(-tX_0)))\}_{t=0} \\
&= (-1)^r [\mathrm{Ad}(g^{-1})(\{X_0, X_r \cdots X_1\})(f \circ \pi)](g) \\
&\quad + (-1)^r [\mathrm{Ad}(g^{-1})(-X_0)(\mathrm{Ad}(g^{-1})(X_r \cdots X_1)(f \circ \pi))](g),
\end{aligned}$$

where the unique derivation of $\boldsymbol{D}(G)$ extending the endomorphism $X \to [X_0, X]$ of $\mathfrak{g}$ has been denoted $D \to \{X_0, D\}$. But it is clear that $\{X_0, D\} = X_0 D - D X_0$ so the two last expressions add up to

$$(-1)^{r+1}[\mathrm{Ad}(g^{-1})(X_r \cdots X_1 X_0)(f \circ \pi)](g).$$

This proves (2). Since G is unimodular, the adjoint X^* equals $-X$ $(X \in \mathfrak{g})$ and the lemma is proved.

COROLLARY 7.3. *$\mu(D^*) = \nu(D)$ for each $D \in \boldsymbol{Z}(G)$.*

In fact $\boldsymbol{Z}(G)$ is invariant under the mapping $D \to D^*$, so $(D^*)^{R(g)} = D^*$.

Remark. The mapping μ of $\boldsymbol{D}_0(G)$ into $\boldsymbol{D}(G/K)$ can be defined for any connected Lie group G and a closed subgroup K, and if G is unimodular, Cor. 7.3 holds.

The image of $\boldsymbol{Z}(G)$ under the mapping μ (or ν) will be denoted $\boldsymbol{Z}(G/K)$.

The Killing form $\langle\ ,\ \rangle$ of $\mathfrak{g}$ induces an isomorphism of the vector space $\mathfrak{g}$ onto its dual $\hat{\mathfrak{g}}$, which then extends to an isomorphism of the symmetric algebra $S(\mathfrak{g})$ onto $S(\hat{\mathfrak{g}})$, the algebra of all polynomial functions on $\mathfrak{g}$. Under this isomorphism the ideal $S(\mathfrak{g})\mathfrak{k}$ goes into the ideal of polynomial functions on $\mathfrak{g}$, which vanish identically on $\mathfrak{p}$. The subalgebra $S(\mathfrak{p}) \subset S(\mathfrak{g})$ goes into an algebra of functions on $\mathfrak{g}$ which under the restriction from $\mathfrak{g}$ to $\mathfrak{p}$, is isomorphic to $S(\hat{\mathfrak{p}})$. Let $I(\mathfrak{g})$ and $I(\hat{\mathfrak{g}})$ denote the set of $\mathrm{Ad}(G)$-invariants in $S(\mathfrak{g})$ and $S(\hat{\mathfrak{g}})$, respectively, $I(\mathfrak{p})$ and $I(\hat{\mathfrak{p}})$ the set of $\mathrm{Ad}_G(K)$-invariants in $S(\mathfrak{p})$ and $S(\hat{\mathfrak{p}})$, respectively.

Let $\mathfrak{u}$ denote the compact real form $\mathfrak{u} = \mathfrak{k} + i\mathfrak{p}$ of the complexification $\mathfrak{g}^c$ of $\mathfrak{g}$. We put $\mathfrak{p}_* = i\mathfrak{p}$ and $U = \mathrm{Int}(\mathfrak{u})$, the adjoint group of $\mathfrak{u}$. Since G has center $\{e\}$ it can be identified with its adjoint group, whereby K is identified with the analytic subgroup of U with Lie algebra $\mathfrak{k}$. Let $I(\hat{\mathfrak{u}})$

denote the ring of $\mathrm{Ad}(U)$-invariant polynomial functions on $\mathfrak{u}$ and $I(\hat{\mathfrak{p}}_*)$ the ring of $\mathrm{Ad}_U(K)$-invariant polynomial functions on $\mathfrak{p}_*$.

The subspace $\mathfrak{a}_* = i\mathfrak{a}$ is a maximal abelian subspace of $\mathfrak{p}_*$. We extend $\mathfrak{a}_*$ to a maximal abelian subalgebra $\mathfrak{t}$ of $\mathfrak{u}$. Let θ denote the involutive automorphism $T + X \rightarrow T - X$ $(T \in \mathfrak{k}, X \in \mathfrak{p}_*)$ of $\mathfrak{u}$. If $H \in \mathfrak{t}$, $\theta H - H \in \mathfrak{p}_*$ and commutes elementwise with $\mathfrak{a}_*$. Hence $\theta\mathfrak{t} \subset \mathfrak{t}$ and $\mathfrak{t} = \mathfrak{t} \cap \mathfrak{k} + \mathfrak{a}_*$. Let $\mathfrak{h} = \mathfrak{t}^c$ $\mathfrak{h}_\mathfrak{p} = \mathfrak{a}^c = \mathfrak{a}_*{}^c$, the superscript c denoting complexification in $\mathfrak{g}^c$. Then $\mathfrak{h}_\mathfrak{p} = \mathfrak{h} \cap \mathfrak{p}^c$. Extending the members of $\mathrm{Ad}(U)$ to $\mathfrak{g}^c$, the Weyl group $W(\mathfrak{h})$ is the group of linear transformations of $\mathfrak{h}$ induced by those members of $\mathrm{Ad}(U)$ which leave $\mathfrak{h}$ invariant, and the Weyl group $W(\mathfrak{h}_\mathfrak{p})$ is the group of linear transformations of $\mathfrak{h}_\mathfrak{p}$ induced by those members of $\mathrm{Ad}_U(K)$ which leave $\mathfrak{h}_\mathfrak{p}$ invariant. Let $I(\hat{\mathfrak{h}})$ and $I(\hat{\mathfrak{h}}_\mathfrak{p})$ denote the corresponding algebras of invariant polynomial functions.

PROPOSITION 7.4. *The following properties of the symmetric space G/K are equivalent:*

(i) $\boldsymbol{Z}(G/K) = \boldsymbol{D}(G/K)$.

(ii) *The restriction from $\mathfrak{g}$ to $\mathfrak{p}$ maps $I(\hat{\mathfrak{g}})$ onto $I(\hat{\mathfrak{p}})$.*

(iii) *The restriction from $\mathfrak{u}$ to $\mathfrak{p}_*$ maps $I(\hat{\mathfrak{u}})$ onto $I(\hat{\mathfrak{p}}_*)$.*

(iv) *The restriction from $\mathfrak{h}$ to $\mathfrak{h}_\mathfrak{p}$ maps $I(\hat{\mathfrak{h}})$ onto $I(\hat{\mathfrak{h}}_\mathfrak{p})$.*

Proof. To prove the equivalence of (i) and (ii) we consider the canonical mapping λ of $S(\mathfrak{g})$ onto $\boldsymbol{D}(G)$ which maps $I(\mathfrak{g})$ onto $\boldsymbol{Z}(G)$ (see [11], p. 192). Let X_i $(1 \leqq i \leqq n)$ be a basis of $\mathfrak{g}$ such that X_i $(1 \leqq i \leqq r)$ is a basis of $\mathfrak{p}$. Then, if $g \in G$, $f \in C^\infty(G)$, $P \in S(\mathfrak{g})$, $\partial_i = \partial/\partial x_i$ we have

$$[\lambda(P)f](g) = [P(\partial_1, \cdots, \partial_n)f(g \exp(x_1X_1 + \cdots + x_nX_n)](0).$$

Now suppose $P \in I(\mathfrak{g})$ and let $\dot{P} \in S(\mathfrak{p})$ be determined by $P - \dot{P} \in S(\mathfrak{g})\mathfrak{k}$. Then $\dot{P} \in I(\mathfrak{p})$. On the other hand there exists a unique $P' \in I(\mathfrak{p})$ such that (see [14], Theorem 10)

$$[\mu(\lambda(P))f](\pi(g)) = [P'(\partial_1, \cdots, \partial_r)(f \circ \pi)(g \exp(x_1X_1 + \cdots + x_rX_r))](0),$$

and $\mu(\lambda(P')) = \mu(\lambda(P))$. It follows that

$$\lambda(\dot{P} - P') + \lambda(P - \dot{P}) = \lambda(P - P') \in \boldsymbol{D}(G)\mathfrak{k}, \tag{3}$$

since μ has kernel $\boldsymbol{D}_0(G) \cap \boldsymbol{D}(G)\mathfrak{k}$. Now suppose P is homogeneous of degree $d > 0$. Then the same holds for $P - \dot{P}$ (0 has any degree) and P' has degree d. But there exists an operator $D_1 \in \boldsymbol{D}(G)$ of order $< d$ such that

$\lambda(P-\bar{P})-D_1 \in \mathbf{D}(G)\mathfrak{k}$. Thus, by (3) and the fact that $\lambda(S(\mathfrak{p})) \cap \mathbf{D}(G)\mathfrak{k} = 0$ we obtain

$$\text{degree}(\bar{P}-P') < \text{degree}(P). \tag{4}$$

The mapping μ maps $\mathbf{Z}(G)$ onto $\mathbf{D}(G/K)$ if and only if the mapping $P \rightarrow P'$ maps $I(\mathfrak{g})$ onto $I(\mathfrak{p})$, which, by (4), happens only if $P \rightarrow \bar{P}$ maps $I(\mathfrak{g})$ onto $I(\mathfrak{p})$. This proves the equivalence of (i) and (ii).

If the elements of $\mathrm{Ad}(G)$ and $\mathrm{Ad}(U)$ are extended to $\mathfrak{g}^c$, a polynomial function on $\mathfrak{g}^c$ is invariant under $\mathrm{Ad}(G)$ if and only if it is invariant under $\mathrm{Ad}(U)$. Since the decompositions $\mathfrak{u} = \mathfrak{k} + \mathfrak{p}_*$ and $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ have a common complexification, namely $\mathfrak{g}^c = \mathfrak{k}^c + \mathfrak{p}^c$, the equivalence of (ii) and (iii) is clear.

Let $I(\hat{\mathfrak{g}}^c)$ and $I(\hat{\mathfrak{p}}^c)$, respectively, denote the algebras of $\mathrm{Ad}(U)$-invariant and $\mathrm{Ad}_U(K)$-invariant polynomial functions on $\mathfrak{g}^c$ and $\mathfrak{p}^c$. By a theorem of Chevalley ([15], p. 430) the restrictions from $\mathfrak{g}^c$ to $\mathfrak{h}$ and from $\mathfrak{p}^c$ to $\mathfrak{h}_{\mathfrak{p}}$ map $I(\hat{\mathfrak{g}}^c)$ isomorphically onto $I(\hat{\mathfrak{h}})$ and $I(\hat{\mathfrak{p}}^c)$ isomorphically onto $I(\hat{\mathfrak{h}}_{\mathfrak{p}})$. The restriction from $\mathfrak{g}^c$ to $\mathfrak{p}^c$ maps $I(\hat{\mathfrak{g}}^c)$ into $I(\hat{\mathfrak{p}}^c)$ so the restriction from $\mathfrak{h}$ to $\mathfrak{h}_{\mathfrak{p}}$ maps $I(\hat{\mathfrak{h}})$ into $I(\hat{\mathfrak{h}}_{\mathfrak{p}})$ and these maps are both onto if and only if one of them is. This proves the equivalence of (iii) and (iv) so Prop. 7.4 is proved.

THEOREM 7.5. (i) *Suppose* $\mathfrak{g}$ *is a classical semisimple Lie algebra (real or complex). Then the restriction from* $\mathfrak{h}$ *to* $\mathfrak{h}_{\mathfrak{p}}$ *maps* $I(\hat{\mathfrak{h}})$ *onto* $I(\hat{\mathfrak{h}}_{\mathfrak{p}})$.

(ii) *Part* (i) *does not hold*[2] *in general for the real forms of the exceptional Lie algebras* $\mathfrak{e}_6$, $\mathfrak{e}_7$, $\mathfrak{e}_8$.

Let Δ denote the set of nonzero roots of $\mathfrak{g}^c$ with respect to $\mathfrak{h}$. Let $\mathfrak{h}_{\mathfrak{p}}{}^*$ denote the annihilator of $\mathfrak{h}_{\mathfrak{p}}$ in the dual space $\hat{\mathfrak{h}}$ of $\mathfrak{h}$, and if f is a function on $\mathfrak{h}$, let $\bar{f}$ denote its restriction to $\mathfrak{h}_{\mathfrak{p}}$.

In the proof below we shall use the well known fact ([21], Theorem 3, [25], p. 108) that $W(\mathfrak{h}_{\mathfrak{p}})$ consist of the restrictions to $\mathfrak{h}_{\mathfrak{p}}$ of those elements in $W(\mathfrak{h})$ which leave $\mathfrak{h}_{\mathfrak{p}}$ invariant.

Proof of (i). It suffices to verify (i) in the case when $\mathfrak{g}$ is a simple Lie algebra over $\mathbf{R}$ and does not have a complex structure. We shall use Cartan's classification of such algebras, listed in [15], Ch. IX. Let $D(a_1,\cdots,a_n)$ denote a diagonal matrix with diagonal elements $a_1,\cdots,a_n$.

If $\mathfrak{g}$ is a normal real form of its complexification, then (i) is obvious.

---

[2] The homomorphisms $\nu\colon \mathbf{Z}(G) \rightarrow D(G/K)$ and $\mu\colon \mathbf{Z}(G) \rightarrow D(G/K)$ are considered in F. A. Berezin [3]. On page 243 it is stated that ν is onto and on page 248 it is stated that μ is onto. Each of these statements is contradicted by Theorems 7.4 and 7.5.

These are the real forms **AI**, **BDI** ($p = q$ and $p + 1 = q$) and **CI**. We shall now verify (i) for the other real simple Lie algebras.

Type AII. Here

$$\begin{aligned} \mathfrak{h}_{\mathfrak{p}} &= \{D(h_1, \cdots, h_n, h_1, \cdots, h_n) : h_i \in \boldsymbol{C}, \sum_1^n h_i = 0\}; \\ \mathfrak{h} &= \{D(h_1, \cdots, h_n, h_{n+1}, \cdots, h_{2n}) : h_i \in \boldsymbol{C}, \sum_1^{2n} h_i = 0\}. \end{aligned} \tag{4}$$

Let e_i $(1 \leqq i \leqq 2n)$ be the linear form on $\mathfrak{h}$ determined by

$$e_i(H) = h_i \qquad \text{if } H = D(h_1, \cdots, h_{2n}). \tag{5}$$

Then the root system Δ is given by

$$\Delta = \{\pm (e_i - e_j) : 1 \leqq i < j \leqq 2n\}.$$

The Weyl group $W(\mathfrak{h})$, acting on the dual of $\mathfrak{h}$, consists of all permutations of the e_i $(1 \leqq i \leqq 2n)$. Hence $I(\hat{\mathfrak{h}})$ consists of those polynomial functions on $\mathfrak{h}$ which can be written as a symmetric function of the e_i. Now if $s \in W(\mathfrak{h})$ leaves $\mathfrak{h}_{\mathfrak{p}}$ invariant its action on $\hat{\mathfrak{h}}$ leaves $\mathfrak{h}_{\mathfrak{p}}^*$ invariant so $s(e_i + e_{i+n}) = e_j + e_{j+n}$. Putting

$$f_i(H) = \tfrac{1}{2}(e_i(H) + e_{i+n}(H)) \qquad (H \in \mathfrak{h}_{\mathfrak{p}}, 1 \leqq i \leqq n)$$

so $f_i = \bar{e}_i = \bar{e}_{i+n}$, one finds that $W(\mathfrak{h}_{\mathfrak{p}})$, acting on $\hat{\mathfrak{h}}_{\mathfrak{p}}$, consists of all permutations of the f_i $(1 \leqq i \leqq n)$. Thus $I(\hat{\mathfrak{h}}_{\mathfrak{p}})$ consists of those polynomial functions on $\mathfrak{h}_{\mathfrak{p}}$ which can be written as a symmetric function of the f_i $(1 \leqq i \leqq n)$. It is now clear that $I(\hat{\mathfrak{h}})^- = I(\hat{\mathfrak{h}}_{\mathfrak{p}})$.

Type AIII. Here we can take $p \leqq q$ and

$$\mathfrak{h}_{\mathfrak{p}} = \sum_1^p \boldsymbol{C}(E_{i\,p+i} + E_{p+i\,i});$$

$$\begin{aligned} \mathfrak{h} = \{H = \sum_1^p h_i(E_{i\,p+i} + E_{p+i\,i}) + \sum_1^p u_i(E_{ii} + E_{p+i\,p+i}) \\ + \sum_{2p+1}^{p+q} u_i E_{ii} : h_i, u_i \in \boldsymbol{C}, \operatorname{Tr}(H) = 0\}. \end{aligned}$$

In fact, this $\mathfrak{h}$ is the image of (4) ($p + q$ for $2n$) under the automorphism $X \to AXA^{-1}$ of $\mathfrak{sl}(p + q, \boldsymbol{C})$, where $A = I - \sum_1^p (E_{i\,p+i} - E_{p+i\,i})$. Hereby, the linear forms (5) correspond to the linear forms

$$e_i(H) = \begin{cases} \tfrac{1}{2}(u_i + h_i) & 1 \leqq i \leqq p, \\ \tfrac{1}{2}(u_i - h_i) & p + 1 \leqq i \leqq 2p, \\ u_i & 2p + 1 \leqq i \leqq p + q. \end{cases}$$

Thus

$$\Delta = \{\pm (e_i - e_j) : 1 \leqq i, j \leqq p + q\},$$

and $W(\mathfrak{h})$, acting on the dual, consists of all permutations of the e_i $(1 \leqq i \leqq p + q)$. Since

$$\mathfrak{h}_{\mathfrak{p}}{}^* = \sum_{2p+1}^{p+q} C e_i$$

it is clear that $\sigma \in W(\mathfrak{h})$ leaves $\mathfrak{h}_{\mathfrak{p}}$ invariant if and only if it permutes the e_i $(2p + 1 \leqq i \leqq p + q)$ and therefore also the e_i $(1 \leqq i \leqq 2p)$. The restrictions $f_i = \bar{e}_i$ $(1 \leqq i \leqq p + q)$ satisfy

$$\bar{e}_i = -\bar{e}_{i+p} \quad (1 \leqq i \leqq p), \qquad \bar{e}_i = 0 \quad (2p + 1 \leqq i \leqq p + q).$$

Therefore $W(\mathfrak{h}_{\mathfrak{p}})$, acting on the dual of $\mathfrak{h}_{\mathfrak{p}}$, consists of all the maps $f_i \rightarrow \epsilon_i f_{\sigma(i)}$, where $\epsilon_i = \pm 1$, and σ is any permutation of $\{1, \cdots, p\}$. Thus $I(\hat{\mathfrak{h}}_{\mathfrak{p}})$ is generated by the polynomials $\sum_1^p f_i^{2k}$ $(1 \leqq k \leqq p)$ which are restrictions of the polynomials $\frac{1}{2} \sum_{i=1}^{p+q} e_i^{2k}$.

Type BDI. Here we use the description in [29], p. 401. The quadratic form corresponding to the symmetric matrix

$$J_{p,q} = \begin{pmatrix} 0 & I_p & 0 \\ I_p & 0 & 0 \\ 0 & 0 & -I_{q-p} \end{pmatrix} \qquad (q \geqq p)$$

has signature (p, q). The corresponding orthogonal group has Lie algebra

$$\mathfrak{g} = \{X \in \mathfrak{gl}(p + q, \boldsymbol{R}) : {}^tXJ_{p,q} + J_{p,q}X = 0\}$$

and a Cartan involution of this Lie algebra is given by $X \rightarrow J_{p,q}XJ_{p,q}$, inducing a Cartan decomposition $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$. Here we can take

$$\mathfrak{h}_{\mathfrak{p}} = \{D(h_1, \cdots, h_p, -h_1, \cdots, -h_p, 0, \cdots, 0) : h_i \in \boldsymbol{C}\}$$

and if $t = [\frac{1}{2}(q - p)]$,

$$\mathfrak{h} = \{H = \sum_1^p h_i(E_{ii} - E_{i+i\,p+p}) + \sum_1^p u_i(E_{2p+t+i\,2p+i} - E_{2p+i\,2p+t+i}) : \ h_i, u_i \in \boldsymbol{C}\}.$$

Let e_i denote the linear form on $\mathfrak{h}$ given by

$$e_j(H) = \begin{cases} h_j & 1 \leqq j \leqq p \\ iu_{j-p} & p + 1 \leqq j \leqq [\frac{1}{2}(p + q)]. \end{cases}$$

Then

$$\Delta = \{\pm(e_i \pm e_j) : 1 \leqq i < j \leqq [\tfrac{1}{2}(p+q)]\}.$$

Now we treat separately the cases $p+q=$ even, $p+q=$ odd.

Type DI, $p+q=2l$. Here $W(\mathfrak{h})$, acting on $\hat{\mathfrak{h}}$, consists of all maps $e_i \to \epsilon_i e_{\sigma(i)}$, where σ is any permutation of $\{1,\cdots,l\}$, $\epsilon_i = \pm 1$, and $\prod_1^l \epsilon_i = 1$. Hence $I(\mathfrak{h})$ is generated by $e_1 \cdots e_l$ and the symmetric polynomials in $e_1^2,\cdots,e_l^2$.

Now if $p=q$, $\mathfrak{g}$ is the normal form of $\mathfrak{g}^c$ so we may assume $p<q$, so $p+1 \leqq l$. Then $s \in W(\mathfrak{h})$ leaves $\mathfrak{h}_\mathfrak{p}$ invariant if and only if it maps $\sum_{p+1}^{l} \boldsymbol{C} e_j$, and therefore also $\sum_1^p \boldsymbol{C} e_j$, into itself. Putting $f_i = \bar{e}_i$ $(1 \leqq i \leqq p)$ we see therefore that $W(\mathfrak{h}_\mathfrak{p})$ consists of all the maps $f_i \to \epsilon_i f_{\sigma(i)}$ where σ is any permutation of $\{1,\cdots,p\}$ and $\epsilon_i = \pm 1$. It is now clear that $(I(\hat{\mathfrak{h}}))^- = I(\hat{\mathfrak{h}}_\mathfrak{p})$.

Type BI, $p+q=2l+1$. Here $W(\mathfrak{h})$ consists of the maps $e_i \to \epsilon_i e_{\sigma(i)}$, ($\sigma$ any permutation of $\{1,\cdots,l\}$, $\epsilon_i = \pm 1$). That $(I(\hat{\mathfrak{h}}))^- = I(\hat{\mathfrak{h}}_\mathfrak{p})$ is seen in the same way as for **DI.**

Type DIII. Using again the description in [29], p. 406, $\mathfrak{g}$ can be regarded as the Lie algebra of the group of linear transformations in $2l$ complex variables leaving invariant

$$z_1 z_{k+1} + \cdots + z_{3k} z_{4k} \; (+ z_{2l-1} z_{2l}),$$
$$z_1 \bar{z}_{2k+1} + \cdots + z_{2k} \bar{z}_{4k} + \bar{z}_1 z_{2k+1} + \cdots + \bar{z}_{2k} z_{4k} \; (+ z_{2l-1} \bar{z}_{2l-1} - z_{2l} \bar{z}_{2l}),$$

where $k = [\tfrac{1}{2} l]$, and the terms in the parentheses occur only if l is odd. The mapping $X \to -{}^t\bar{X}$ is a Cartan involution of $\mathfrak{g}$, inducing a Cartan decomposition $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ of $\mathfrak{g}$. Here we can take (cf. [29], p. 408)

$$\begin{aligned}\mathfrak{h}_\mathfrak{p} = \{D(h_1,\cdots,h_k,-h_1,\cdots,-h_k,-h_1,\cdots,\\ -h_k,h_1,\cdots,h_k,(0,0)) : h_i \in \boldsymbol{C}\},\end{aligned}$$

$$\begin{aligned}\mathfrak{h} = \{H = D(h_1,\cdots,h_k,-h_1,\cdots,-h_k,h_{k+1},\cdots,\\ h_{2k},-h_{k+1},\cdots,-h_{2k},(h_{2k+1},-h_{k+1})) : h_i \in \boldsymbol{C}\},\end{aligned}$$

the terms in the parentheses occurring only if n is odd. If e_i is the linear form on $\mathfrak{h}$ given by

$$e_i(H) = h_i \qquad (1 \leqq i \leqq l)$$

then

$$\Delta = \{\pm(e_i \pm e_j) : 1 \leqq i < j \leqq l\}.$$

Here $W(\mathfrak{h})$ consists of all the maps $e_i \to \epsilon_i e_{\sigma(i)}$, where σ is any permutation of $\{1, \cdots, l\}$, $\epsilon_i = \pm 1$ and $\prod_1^l \epsilon_i = 1$. Hence $I(\mathfrak{h})$ is the same as for Type **DI**. Now $\mathfrak{h}_{\mathfrak{p}}^* = \sum_{i=1}^{k} \boldsymbol{C}(e_i + e_{i+k}) + (\boldsymbol{C}e_{2k+1})$; if $s \in W(\mathfrak{h})$ leaves $\mathfrak{h}_{\mathfrak{p}}$ invariant, it leaves $\mathfrak{h}_{\mathfrak{p}}^*$ invariant and consequently $s(e_j + e_{j+k}) = \pm (e_i + e_{i+k})$. Hence $s(e_j - e_{j+k}) = \pm (e_i - e_{i+k})$. If we put $f_j = \bar{e}_j - \bar{e}_{j+k}$ $(1 \leqq j \leqq k)$ we find therefore that the restriction of s to $\mathfrak{h}_{\mathfrak{p}}$, acting on $\hat{\mathfrak{h}}_{\mathfrak{p}}$, is of the form $f_j \to \epsilon_j f_{\sigma(j)}$ where $\epsilon_j = \pm 1$ and σ is a permutation of $\{1, \cdots, k\}$. It is also clear that each such σ occurs in this manner; the ϵ_j are also unrestricted since a change of sign of some f_j is induced by a change of sign of an even number of e_i. The transformations $f_j \to \epsilon_j f_{\sigma(j)}$ $(\epsilon_j = \pm 1$, σ arbitrary permutation of $\{1, \cdots, k\})$ which are all different, the f_j being linearly independent, exhaust the group $W(\mathfrak{h}_{\mathfrak{p}})$; the algebra $I(\mathfrak{h}_{\mathfrak{p}})$ is generated by $\sum_{i=1}^{k} f_i^{2n}$ $(n = 1, \cdots, k)$. Also, if $H \in \mathfrak{h}_{\mathfrak{p}}$

$$\bar{e}_i(H) = h_i \quad (1 \leqq i \leqq k), \qquad \bar{e}_i(H) = -h_{i-k} \quad (k + 1 \leqq i \leqq 2k),$$

so

$$\sum_{i=1}^{l} (\bar{e}_i)^{2n} = 2 \sum_{i=1}^{k} f_i^{2n},$$

proving $(I(\hat{\mathfrak{h}}))^- = I(\hat{\mathfrak{h}}_{\mathfrak{p}})$.

Type CII. Here $\mathfrak{g} = \mathfrak{sp}(p, q)$ and we suppose $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ is the Cartan decomposition given in [15], p. 351. We assume $p \leqq q$. The complexification $\mathfrak{g}^c = \mathfrak{sp}(p + q, \boldsymbol{C})$ has a Cartan subalgebra

$$\mathfrak{h}' = \{H' = D(h_1, \cdots, h_{p+q}, -h_1, \cdots, -h_{p+q}) : h_i \in \boldsymbol{C}\}.$$

Let J denote the matrix in $\boldsymbol{Sp}(\boldsymbol{n}, \boldsymbol{C})$ given by

$$J = 2^{-\frac{1}{2}}\{I + \sum_1^p (E_{i\,p+i} - E_{p+i\,i}) + \sum_1^p (E_{p+q+i\;2p+q+i} - E_{2p+q+i\;p+q+i})\}$$

and let $\mathfrak{h}$ denote the image of $\mathfrak{h}'$ under the automorphism $X \to JXJ^{-1}$ of $\mathfrak{sp}(p + q, \boldsymbol{C})$. We have

$$H = JH'J^{-1} = D(u_1, \cdots, u_{p+q}, -u_1, \cdots, -u_{p+q}) + \sum_1^p k_i(E_{i\,p+i} + E_{p+i\,i})$$
$$- \sum_1^p k_i (E_{p+q+i\;2p+q+i} + \; + E_{2p+q+i\;p+q+i})$$

where

$$u_i = \tfrac{1}{2}(h_i + h_{i+p}), \quad (1 \leqq i \leqq p), \qquad u_j = \tfrac{1}{2}h_j, \quad (2p + 1 \leqq j \leqq p + q)$$
$$k_i = \tfrac{1}{2}(-h_i + h_{i+p}), \quad (1 \leqq i \leqq p).$$

The intersection $\mathfrak{h} \cap \mathfrak{p}^c$ is

$$\sum_1^p C(E_{i\,p+i} + E_{p+i\,i} - E_{p+q+i\;2p+q+i} - E_{2p+q+i\;p+q+i}),$$

which can serve as $\mathfrak{h}_\mathfrak{p}$. Now consider the linear forms e_i' on $\mathfrak{h}'$ defined by

$$e_i'(H) = h_i \quad (1 \leqq i \leqq p+q).$$

Then if Δ' denotes the set of nonzero roots of $\mathfrak{g}^c$ with respect to $\mathfrak{h}'$,

$$\Delta' = \{\pm (e_i' \pm e_j'), 1 \leqq i < j \leqq p+q; \pm 2e_i' \ (1 \leqq i \leqq p+q)\}.$$

The Weyl group $W(\mathfrak{h}')$ consists of all maps $e_i' \to \epsilon_i e'_{\sigma(i)}$, where $\epsilon_i = \pm 1$ and σ is an arbitrary permutation of $\{1, 2, \cdots, n\}$. Consider the linear form e_i on $\mathfrak{h}$ defined by $e_i(JH'J^{-1}) = e_i'(H')$, $H' \in \mathfrak{h}'$. Then

$$e_i(H) = \begin{cases} u_i - k_i & 1 \leqq i \leqq p \\ u_{i-p} + k_{i-p} & p+1 \leqq i \leqq 2p \\ 2u_i & 2p+1 \leqq i \leqq p+q \end{cases}$$

and

$$\Delta = \{\pm (e_i \pm e_j), 1 \leqq i < j \leqq p+q; \pm 2e_i \ (1 \leqq i \leqq p+q)\},$$

and $W(\mathfrak{h})$ consists of all maps $e_i \to \epsilon_i e_{\sigma(i)}$, where $\epsilon_i = \pm 1$ and σ arbitrary permutation of $\{1, 2, \cdots, n\}$. Also

$$\mathfrak{h}_\mathfrak{p}^* = \sum_1^p C(e_i + e_{i+p}) + \sum_{2p+1}^{p+q} Ce_i. \tag{6}$$

If $s \in W(\mathfrak{h})$ leaves $\mathfrak{h}_\mathfrak{p}^*$ invariant it is easily seen that s, acting on $\hat{\mathfrak{h}}$, leaves each of the sums in (6) invariant. Consequently $s(e_i - e_{i+p}) = \pm (e_j - e_{j+p})$. If we put $f_j = \bar{e}_j - \bar{e}_{j+p}$ $(1 \leqq j \leqq p)$, the restriction of s to $\mathfrak{h}_\mathfrak{p}$, acting on $\hat{\mathfrak{h}}_\mathfrak{p}$, is of the form $f_j \to \epsilon_j f_{\sigma(j)}$ ($\epsilon_j = \pm 1$, σ any permutation of $\{1, \cdots, p\}$ and any such transformation of $\hat{\mathfrak{h}}_\mathfrak{p}$ arises from some $s \in W(\mathfrak{h})$. The f_j $(1 \leqq j \leqq p)$ being linearly independent, these transformations exhaust $W(\mathfrak{h}_\mathfrak{p})$; the algebra $I(\hat{\mathfrak{h}}_\mathfrak{p})$ is generated by $\sum_1^p f_j^{2n}$ $(n = 1, \cdots, p)$ and $I(\hat{\mathfrak{h}})$ is generated by $\sum_1^{p+q} e_i^{2n}$ $(n = 1, \cdots, p+q)$. Since

$$\bar{e}_i = \begin{cases} \tfrac{1}{2} f_i & (1 \leqq i \leqq p), \\ \tfrac{1}{2} f_{i-p} & (p+1 \leqq i \leqq 2p), \\ 0 & (2p+1 \leqq i \leqq p+q), \end{cases}$$

it is clear that $I(\hat{\mathfrak{h}})^- = I(\hat{\mathfrak{h}}_\mathfrak{p})$.

Proof of (ii). Consider first the noncompact symmetric space of type E IV, that is E_6/F_4. Here $I(\hat{\mathfrak{h}})$ is generated by algebraically independent generators of degrees 2, 5, 6, 8, 9, 12 ([31], [5]). For this symmetric space the set Σ_0 from § 2 is the set of positive roots for the Lie algebra of $\boldsymbol{SU}(3)$ (see [4], p. 422 or [1], p. 33). It follows that $W(\mathfrak{h}_\mathfrak{p})$ is the Weyl group of $\boldsymbol{SU}(3)$ so $I(\hat{\mathfrak{h}}_\mathfrak{p})$ has generators i_2, i_3 of degrees 2 and 3, respectively. However $I(\hat{\mathfrak{h}})$ cannot contain any polynomial of degree 3 so $I(\hat{\mathfrak{h}})^- \neq I(\hat{\mathfrak{h}}_\mathfrak{p})$. For the noncompact symmetric space of type E VII, that is $E_7/E_6\cdot\boldsymbol{T}$ ($\boldsymbol{T} =$ 1-torus) we have $I(\hat{\mathfrak{h}})$ generated by algebraically independent generators of degrees 2, 6, 8, 10, 12, 14, 18. Here the set Σ_0 is the set of positive roots for the Lie algebra of $\boldsymbol{Sp}(3)$ (see [4], p. 423 or [1], p. 33). Thus $W(\mathfrak{h}_\mathfrak{p})$ is the Weyl group of $\boldsymbol{Sp}(3)$ and $I(\hat{\mathfrak{h}}_\mathfrak{p})$ is generated by algebraically independent generators i_2, i_4, i_6 of degrees 2, 4, 6 respectively. By the algebraic independence, $i_4 \notin I(\hat{\mathfrak{h}})^-$. For the noncompact symmetric space of type E IX, that is $E_8/E_7\cdot\boldsymbol{SU}(2)$, we have $I(\hat{\mathfrak{h}})$ generated by algebraically independent generators of degrees 2, 8, 12, 14, 18, 20, 24, 30. The set Σ_0 is now the set of positive roots for the Lie algebra of F_4 so $I(\hat{\mathfrak{h}}_\mathfrak{p})$ has algebraically independent generators of degrees 2, 6, 8, 12. The one of degree 6 cannot belong to $I(\hat{\mathfrak{h}})^-$.

This concludes the proof of Theorem 7. 5.

We observe now that in each of the three examples above the generators of $I(\hat{\mathfrak{h}}_\mathfrak{p})$ which do not lie in $(I(\hat{\mathfrak{h}}))^-$ have degrees $d - e$ where d and e are degrees of certain generators of $I(\hat{\mathfrak{h}})$. This raises the question whether the elements in $I(\hat{\mathfrak{h}}_\mathfrak{p})$ are all obtained from rational invariants on $\mathfrak{h}$ by restriction. We shall now prove that this is indeed the case.

THEOREM 7. 6. *Let $\boldsymbol{C}(I(\hat{\mathfrak{h}}))$ and $\boldsymbol{C}(I(\hat{\mathfrak{h}}_\mathfrak{p}))$ denote the quotient fields of $I(\hat{\mathfrak{h}})$ and $I(\hat{\mathfrak{h}}_\mathfrak{p})$, respectively. Then the restriction $f \to \bar{f}$ from $\mathfrak{h}$ to $\mathfrak{h}_\mathfrak{p}$ induces a mapping of $\boldsymbol{C}(I(\hat{\mathfrak{h}}))$ onto $\boldsymbol{C}(I(\hat{\mathfrak{h}}_\mathfrak{p}))$.*

Let $U = \mathrm{Int}(\mathfrak{u})$ as before and let θ denote the automorphism $T + X \to T - X$ ($T \in \mathfrak{t}, X \in \mathfrak{p}_*$) of $\mathfrak{u}$. The corresponding automorphism of U will also be denoted by θ, and K_* shall denote the group of fixed points of θ. Let A_* denote the analytic subgroup $\exp(\mathfrak{a}_*)$ of U and put $P_* = \exp(\mathfrak{p}_*)$.

LEMMA 7. 7. *A_* equals its centralizer in P_*.*

Proof. Suppose $p \in P_*$ commutes elementwise with A_*. Let B' denote the subgroup of U generated by p and A_*, B the closure of B' in U. Clearly, B consists of all products ap^n ($a \in A_*$, $n =$ integer). Now $a = b^2$ ($b \in A_*$) and $p = q^2$ ($q \in P_*$) so $ap^n = b^2p^n = bq^nq^nb = bq^n\theta((bq^n)^{-1})$. Since $P^* = \{u\theta(u^{-1}) : u \in U\}$ it follows that $B' \subset P_*$; but P_* being closed we have

$B \subset P_*$. The identity component B_0 of B contains some power of p, so B_0 is a torus and B/B_0 is cyclic. As is well known this implies that there exists an element $b \in B$ such that the sequence $e, b, b^2, \cdots$ is dense in B. Select $X \in \mathfrak{p}_*$ such that $b = \exp X$. Then the closure of the one-parameter subgroup $\{\exp tX : t \in \boldsymbol{R}\}$ is a torus, contained in P_*, but containing p and A_*. But $\mathfrak{a}_*$ is maximal abelian in $\mathfrak{p}_*$ so this torus equals A_*, whence $p \in A_*$.

LEMMA 7.8. *Let M' denote the normalizer of $\mathfrak{a}_*$ in K_*. Then $M'A_*$ is the normalizer of $\mathfrak{a}_*$ in U.*

Proof. Suppose $u \in U$ such that $\mathrm{Ad}(u)\mathfrak{a}_* \subset \mathfrak{a}_*$. Then $u = kp$, $(k \in K_*$, $p = \exp X, X \in \mathfrak{p}_*)$ so we conclude $e^{\mathrm{ad}\, X}(\mathfrak{a}_*) \subset \mathfrak{p}_*$, and therefore

$$(\sinh(\mathrm{ad}\, X))(H) = 0 \qquad\qquad \text{for } H \in \mathfrak{h}_{\mathfrak{p}}.$$

Now $\mathrm{ad}\, X$, acting on the complexification $\mathfrak{u}^c$ of $\mathfrak{u}$, is completely reducible and has purely imaginary eigenvalues. Let $e_1, \cdots, e_n$ be eigenvectors of $\mathrm{ad}\, X$ spanning $\mathfrak{u}^c$ and $\lambda_1, \cdots, \lambda_n$ the corresponding eigenvalues. Let

$$H = \sum_1^n a_j(H) e_j \qquad\qquad (H \in \mathfrak{h}_{\mathfrak{p}}).$$

Suppose $e_1, \cdots, e_m$ are all the eigenvectors for which $a_j \not\equiv 0$. Since

$$0 = \sinh(\mathrm{ad}\, X)(H) = \sum_1^m (\sinh \lambda_j) a_j(H) e_j$$

it follows that each λ_j $(1 \leqq j \leqq m)$ is an integral multiple of πi. Consequently

$$e^{\mathrm{ad}(2X)}(H) = H \qquad\qquad H \in \mathfrak{h}_{\mathfrak{p}}$$

so p^2 commutes elementwise with A_*. By Lemma 7.7, $p^2 = \exp 2H_0$ for some $H_0 \in \mathfrak{a}_*$. Put $p_0 = \exp H_0$. Then $\theta(p_0 p^{-1}) = p_0^{-1} p = p_0 p^{-1}$ so $p = k_0 p_0$ where $k_0 \in K_*$. Thus $u = k' \exp H_0$ $(k' \in K_*)$ whence $k' \in M'$.

LEMMA 7.9. *Let J denote the set of restrictions of the polynomials in $I(\hat{\mathfrak{h}})$ to $\mathfrak{h}_{\mathfrak{p}}$. Let $S(\hat{\mathfrak{h}}_{\mathfrak{p}})$ denote the symmetric algebra over $\hat{\mathfrak{h}}_{\mathfrak{p}}$. Then $S(\hat{\mathfrak{h}}_{\mathfrak{p}})$ is integral over J.*

Proof. Let $p \in S(\hat{\mathfrak{h}})$, the symmetric algebra over $\hat{\mathfrak{h}}$. Then p satisfies the equation $\prod_{s \in W(\mathfrak{h})} (x - p^s) = 0$ with leading coefficients 1 and coefficients in $I(\hat{\mathfrak{h}})$. This means that the ring $S(\hat{\mathfrak{h}})$ is integral over $I(\hat{\mathfrak{h}})$. The restriction $p \to \bar{p}$ gives a homomorphism of $S(\hat{\mathfrak{h}})$ onto $S(\hat{\mathfrak{h}}_{\mathfrak{p}})$ mapping $I(\hat{\mathfrak{h}})$ onto J. Hence the lemma.

Passing now to the proof of Theorem 7.6, suppose the quotient field $\boldsymbol{C}(J)$ were a proper subset of $\boldsymbol{C}(I(\hat{\mathfrak{h}}_\mathfrak{p}))$. Then there exists $\alpha \in I(\hat{\mathfrak{h}}_\mathfrak{p})$ such that $\alpha \notin \boldsymbol{C}(J)$. By Lemma 7.9, α is algebraic over $\boldsymbol{C}(J)$. Since $I(\hat{\mathfrak{h}})$ is finitely generated, the homomorphic image J also has finitely many generators, say $\xi_1, \cdots, \xi_r$. Let

$$p_\alpha(x) = x^n + f_1(\xi)x^{n-1} + \cdots + f_n(\xi) \qquad (n \geqq 2)$$

be the polynomial with coefficients in the field $\boldsymbol{C}(J) = \boldsymbol{C}(\xi_1, \cdots, \xi_r)$ of lowest degree having α as a zero and leading coefficient 1. Let $f(\xi)$ denote the product of the denominators of the $f_i(\xi)$ and let $q(\xi)$ denote the product of $f(\xi)$ with the discriminant of the polynomial $f(\xi)p_\alpha(x)$. Then $q(\xi)$ is a polynomial function on $\mathfrak{h}_\mathfrak{p}$, not identically 0. Select $H_* \in \mathfrak{a}_*$ such that

$$(7) \qquad q(\xi)(H_*) \neq 0, \{\exp tH_* \colon t \in \boldsymbol{R} \text{ dense in } A_*\}$$

Consider the homomorphism $\lambda\colon p \to p(H_*)$ of J into $\boldsymbol{C}$. The image of $p_\alpha(x)$ under λ is a polynomial with discriminant $\neq 0$, which therefore has n distinct roots, say $\alpha_1, \cdots, \alpha_n \in \boldsymbol{C}$. For each fixed i $(1 \leqq i \leqq n)$ we wish to extend λ to a homomorphism $\lambda_i\colon \boldsymbol{C}[\xi_1, \cdots, \xi_r, \alpha] \to \boldsymbol{C}$ by putting $\lambda_i(\alpha) = \alpha_i$. To see that this is possible, suppose $p(\xi, x)$ is a polynomial with coefficients in $J = \boldsymbol{C}[\xi_1, \cdots, \xi_r]$ which has α as a zero. Then $p_\alpha(x)$ divides $p(\xi, x)$, i.e. $p(\xi, x) = p_\alpha(x)q(\xi, x)$ and the coefficient of $q(\xi, x)$ are rational expressions in $\xi_1, \cdots, \xi_r$ whose denominators divide $f(\xi)$. Since $\lambda(f(\xi)) \neq 0$ the image of $p(\xi, x)$ under λ has $x = \alpha_i$ as a zero. Thus λ_i is well defined. Since $S(\hat{\mathfrak{h}}_\mathfrak{p})$ is integral over J, and *a fortiori* over $\mathbf{C}[\xi_1, \cdots, \xi_r, \alpha]$, there exists a homomorphism $\Lambda_i\colon S(\hat{\mathfrak{h}}_\mathfrak{p}) \to \boldsymbol{C}$ extending λ_i. Then there exists $H_i \in \mathfrak{h}_\mathfrak{p}$ such that $\Lambda_i(p) = p(H_i)$ for all $p \in S(\hat{\mathfrak{h}}_\mathfrak{p})$. In particular

$$p(H_*) = p(H_1) = p(H_2) \qquad \text{for } p \in I(\mathfrak{h}),$$

from which it follows easily that H_*, H_1, H_2 are all conjugate under $W(\mathfrak{h})$. Since $H_* \in \mathfrak{a}_*$, and since $\mathfrak{t}$ is invariant under $W(\mathfrak{h})$, we obtain $H_1, H_2 \in \mathfrak{t} \cap \mathfrak{h}_\mathfrak{p} = \mathfrak{a}_*$. Now select $u_1, u_2 \in U$ such that $\operatorname{Ad}(u_i)H_* = H_i$ $(i = 1, 2)$. Then $u_i \exp tH_* u_i^{-1} = \exp tH_i$ $(i = 1, 2)$ so by (7) u_1 and u_2 belong to the normalizer of $\mathfrak{a}_*$ in U. Using Lemma 7.8 we have $\operatorname{Ad}(m)H_1 = H_2$ for some $m \in M'$. But the restrictions of $\operatorname{Ad}(M')$ to $\mathfrak{a}_*$ constitute the Weyl group $W(\mathfrak{a}_*)$ although K_* may be larger than K (see [15], p. 244 and p. 273). Hence $sH_1 = H_2$ for some $s \in W(\mathfrak{a}_*)$. But then

$$\alpha_1 = \Lambda_1(\alpha) = \alpha(H_1) = \alpha(H_2) = \Lambda_2(\alpha) = \alpha_2,$$

which is a contradiction.

Remarks. (i) Theorem 7.6 had been proved independently by Harish-Chandra (cf. also his proof in [15], p. 433 of the quoted result of Chevalley). (ii) (*Added in proof*). Since $W(\mathfrak{h}_\mathfrak{p})$ is the group of linear transformations of $\mathfrak{h}_\mathfrak{p}$ induced by those elements of $W(\mathfrak{h})$ which leave $\mathfrak{h}_\mathfrak{p}$ invariant, Theorem 7.6 raises the following more general question: Let V be a vector space over C and G a finite group of linear transformations of V. Let W be a vector subspace of V and H the group of linear transformations of W induced by those members of G which leave W invariant. Does the restriction from V to W map the set of rational G-invariants on V onto the set of rational H-invariants on V onto the set of rational H-invariants on W? According to a communication from D. Mumford, the answer is affirmative.

REFERENCES.

[1] S. Araki, "On root systems and an infinitesimal classification of irreducible symmetric spaces," *J. Math. Osaka City Univ.*, vol. 13 (1962), pp. 1-34.

[2] L. Ásgeirsson, "Some hints on Huygens' principle and Hadamard's conjecture," *Communications on Pure and Applied Mathematics*, vol. 9 (1956), pp. 307-326.

[3] F. A. Berezin, "Laplace operators on semisimple Lie groups," *Trudy Moscovskogo Matematičeskogo Obščestva*, vol. 6 (1957), pp. 371-463 or *American Mathematical Society Translations*, vol. 21, pp. 239-339.

[4] É. Cartan, "Sur certaines formes riemanniennes remarquables des géométries a groupe fondamental simple," *Annales Scientifiques de l'École Normale Supérieure*, vol. 44 (1927), pp. 345-467.

[5] C. Chevalley, "The Betti numbers of the exceptional simple Lie groups," *Proceedings of the International Congress of Mathematicians*, Harvard II (1950), pp. 21-24.

[6] R. Courant and D. Hilbert, *Methoden der mathematischen Physik*, II, Springer, Berlin, 1937.

[7] E. B. Dynkin, "Non-negative eigenfunctions of the Laplace-Beltrami operator and Brownian motion in certain symmetric spaces," *Doklady Akademii Nauk SSSR*, vol. 141 (1961), pp. 288-291; *Soviet Mathematics*, vol. 2 (1961), pp. 1433-1436.

[8] S. G. Gindikin and F. I. Karpelevič, "Plancherel measure of Riemannian symmetric spaces of nonpositive curvature," *Doklady Akademii Nauk SSSR*, vol. 145 (1962), pp. 252-255; *Soviet Mathematics*, vol. 3 (1962), pp. 962-965.

[9] R. Godement, "A theory of spherical functions, I," *Transactions of the American Mathematical Society*, vol. 73 (1952), pp. 496-556.

[10] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations*, Yale Univ. Press, 1923.

[11] Harish-Chandra, "Representations of semisimple Lie groups I," *Transactions of the American Mathematical Society*, vol. 75 (1953), pp. 185-243.

[12] ———, "Differential operators on a semisimple Lie algebra," *American Journal of Mathematics*, vol. 79 (1957), pp. 87-120.

[13] ———, "Spherical functions on a semisimple Lie group I, II," *American Journal of Mathematics*, vol. 80 (1958), pp. 241-310, 553-613.

[14] S. Helgason, "Differential operators on homogeneous spaces," *Acta Mathematica*, vol. 102 (1959). pp. 239-299.

[15] ———, *Differential geometry and symmetric spaces*, Academic Press, New York, 1962.

[16] ———, "Some results on invariant theory," *Bulletin of the American Mathematical Society*, vol. 68 (1962), pp. 367-371.

[17] ———, "Duality and Radon transform for symmetric spaces," *American Journal of Mathematics*, vol. 85 (1963), pp. 667-692.

[18] ———, "Fundamental solutions for invariant differential operators on symmetric spaces," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 778-781.

[19] L. Hörmander, "On the division of distributions by polynomials," *Arkiv för Matematik*, vol. 3 (1958), pp. 555-568.

[20] ———, *Linear partial differential operators*, Springer, Berlin, 1963.

[21] B. Kostant, "On the conjugacy of real Cartan subalgebras I," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 41 (1955), pp. 967-970.

[22] R. P. Langlands, "The dimension of spaces of automorphic forms," *American Journal of Mathematics*, vol. 85 (1963), pp. 99-125.

[23] S. Łojasiewicz, "Division d'une distribution par une fonction analytique de variables reelles," *Comptes Rendus des Séances de l'Académie des Sciences*, Paris, vol. 246 (1958), pp. 683-686.

[24] G. de Rham, *Variétés différentiables*, Hermann, Paris, 1955.

[25] I. Satake, "On representations and compactifications of symmetric Riemann spaces," *Annals of Mathematics*, vol. 71 (1960), pp. 77-110.

[26] K. Stellmacher, "Ein Beispiel einer Huygensschen Differentialgleichung," *Göttingen Nachrichten*, 1953, pp. 133-138.

[27] L. Schwartz, *Théorie des Distributions II*, Hermann, Paris, 1951.

[28] ———, *Séminaire* 1954-55 (Equations aux dérivées partielles), Paris, 1955.

[29] M. Suguira, "Conjugate classes of Cartan subalgebras in real semisimple Lie algebras," *Journal of the Mathematical Society of Japan*, vol. 11 (1959), pp. 374-434.

[30] E. C. Titchmarsh, *The theory of functions*, 2nd ed., Oxford Univ. Press, London, 1939.

[31] Chih-Ta Yen, "Sur les polynomes de Poincaré des groupes de Lie exceptionnels," *Comptes Rendus des Séances de l'Académie des Sciences*, Paris, vol. 228 (1949), pp. 628-630.

COMPACTIFICATIONS AND SEMI-NORMAL SPACES.

By Orrin Frink.

The completely regular T_1 spaces or Tychonoff spaces are those with Hausdorff compactifications. They form a natural generalization of the metric spaces, they may be provided with a uniform structure, and may be characterized in terms of notions of proximity and separation [5]. However, no neat internal characterization of Tychonoff spaces seems to be known. In this note we provide such a characterization in terms of the notion of a *normal base* for the closed sets of a space. A normal base Z for the closed sets of a space X is a base which is a disjunctive ring of sets, disjoint members of which may be separated by disjoint complements of members of Z. The ring of all closed sets is a normal base if X is normal. We call a T_1 space having at least one normal base *semi-normal.* Unlike normality, the property of being semi-normal is hereditary and productive.

In a Tychonoff space X, the zero-sets of real continuous functions over X form a normal base. Hence every Tychonoff space is semi-normal. More generally, if Y is a Hausdorff compactification of X, and R is a lattice-ordered ring of real continuous functions over X which distinguishes points, and is such that the functions of R may be extended to continuous functions over Y, then the zero-sets of members of R form a normal base for X. Conversely it can be shown that if Z is a normal base, then the Wallman space $w(Z)$ consisting of the Z-ultrafilters, is a Hausdorff compactification of X. It follows that every semi-normal space is Tychonoff.

By choosing different normal bases Z for a non-compact semi-normal space X, different Hausdorff compactifications of X may be obtained in the form of Wallman spaces $w(Z)$. It is natural to ask whether every Hausdorff compactification may be obtained in this way. I have not yet been able to answer this question.

It is also natural to ask which real functions over a Tychonoff space X may be extended to continuous real functions over the compactification $w(Z)$ associated with a normal base Z for X. We answer this question by showing that the functions which can be extended are those which are *Z-uniformly continuous.* We give also a similar criterion for the possibility of extending a continuous map of X into a compact Hausdorff space H to a continuous map of $w(Z)$ into H.

Received November 27, 1963.
Revised March 30, 1964.

Definitions. A base Z for the closed sets of a T_1 space X is said to be *disjunctive* if, given any closed set F and any point x not in F, there exists a closed set A of Z which contains x and is disjoint from F. This is similar to the disjunction property of Wallman.

A family of sets is called a *ring of sets* if it contains all finite unions and intersections of its members. Every ring of sets is a lattice.

A base Z for the closed sets of a T_1 space is called a *normal base* if it is a disjunctive ring of closed sets such that any two disjoint members A and B of Z are subsets respectively of disjoint complements C' and D' of members C and D of Z; that is, $A \subset C'$, $B \subset D'$, and $C' \cap D' = \emptyset$.

Members of a normal base Z will be called *Z-sets,* and their complements *Z-complements.* The Z-complements form a base for the open sets of the space.

A proper subset of a normal base Z is called a *Z-filter* if it is closed under finite intersection and contains every superset in Z of each of its members. No Z-filter contains the empty set $\emptyset$.

A *Z-ultrafilter* is a maximal Z-filter. It follows from Zorn's lemma that every Z-filter is contained in at least one Z-ultrafilter.

A T_1 space X is called *semi-normal* if there exists at least one normal base Z for the closed sets of X.

THEOREM 1. *A T_1 space is completely regular if and only if it is semi-normal.*

Proof. If the space X is completely regular, it is well known, and easily verified, that the family Z of all zero-sets of X, that is, of sets $f^{-1}(0)$ where continuous real functions $f(x)$ defined over X take the value zero, is a disjunctive base for the closed sets of X. It is shown in the book "Rings of Continuous Functions" by L. Gillman and M. Jerison [3], that Z is a ring of sets, and that disjoint Z-sets may be separated by disjoint Z-complements. Hence Z is a normal base for X. It follows that every completely regular T_1 space is semi-normal.

Conversely, if X is a semi-normal space with normal base Z, we wish to show that X is homeomorphic to a dense subset of a certain compact Hausdorff space $w(Z)$, namely the Wallman space consisting of Z-ultrafilters. It will follow that X is T_1 and completely regular.

We define a topology in the set $w(Z)$ consisting of all Z-ultrafilters in the usual way, by taking as a base for the closed sets of $w(Z)$ the family of all sets A^*, where A^* is the set of all Z-ultrafilters having A as a member, A being an arbitrary member of Z. This is equivalent to taking as a base for the open sets of $w(Z)$ the family of all sets U^* consisting of all Z-ultrafilters having at least one subset of U as a member, U being an arbitrary Z-complement.

The space X is imbedded in the space $w(Z)$ by the imbedding map, which sends a point x of X into the Z-ultrafilter consisting of all Z-sets which contain x. This imbedding maps X into a dense subset of $w(Z)$, since every non-empty basic open set of type U^* contains the image of at least one point of X. That the imbedding map is one-to-one follows from the fact that the base Z is disjunctive. That the imbedding map is a homeomorphism follows from the correspondence between the basic open sets U and U^*.

That the space $w(Z)$ is compact may be seen as follows. Corresponding to any family having the finite intersection property, and consisting of basic closed sets of type A^* in $w(Z)$, is the family consisting of the corresponding Z-sets A in X, which has also the finite intersection property. The extension of this family of Z-sets to be maximal is a Z-ultrafilter, which is seen to lie in every basic set A^* of the original family. Hence this family has a non-empty intersection. Since the sets A^* form a base for the closed sets of $w(Z)$, it follows that $w(Z)$ is compact.

Finally it may be seen that $w(Z)$ is a Hausdorff space. Any two distinct Z-ultrafilters $\mathfrak{A}$ and $\mathfrak{B}$ must contain disjoint Z-sets A and B. Since Z is a normal base, A and B must lie respectively in disjoint open Z-complements U and V. Hence the disjoint open sets U^* and V^* in $w(Z)$ are open sets containing $\mathfrak{A}$ and $\mathfrak{B}$, and the space has the Hausdorff separation property.

The foregoing proof that $w(Z)$ is a Hausdorff compactification of X is similar to Wallman's original proof for the case where X is normal and Z is the family of all closed sets of X [9]. It is similar also to Kelley's proof of Wallman's theorem [4]. It uses some of the terminology of Gillman and Jerison, who show in [3] that if Z is the family of all zero-sets of real continuous functions over a Tychonoff space X, then $w(Z)$ is homeomorphic to the Stone-Čech compactification βX. We have modified the Gillman-Jerison procedure by requiring only that the normal base Z have those properties of the family of zero-sets which are needed to carry through the proof that $w(Z)$ is a Hausdorff compactification of X.

Continuous extensions. It is natural to ask what bounded continuous real functions on a space X may be extended to functions which are continuous on a particular Hausdorff compactification of X, such as $w(Z)$. Many results of this general nature are well known. For example, if X is a subspace of a metric space Y, the real functions on X that may be extended to continuous functions on Y are those that are uniformly continuous relative to the metric of Y. There are similar theorems for uniform spaces. We define here a related condition called Z-uniform continuity for spaces with a normal base Z.

Definition. A real function $f(x)$ defined over a semi-normal space X with normal base Z is said to be *Z-uniformly continuous* if for every positive epsilon there exists a finite open cover of X by Z-complements, on each of which the oscillation of $f(x)$ is less than epsilon.

THEOREM 2. *A real function $f(x)$ defined over a semi-normal space X with normal base Z can be extended to a real continuous function over the compactification $w(Z)$ if and only if $f(x)$ is Z-uniformly continuous.*

Proof. The condition is necessary, for suppose $f(y)$ is a continuous real function defined over the compact space $w(Z)$. Given a positive epsilon, we may cover $w(Z)$ by a finite number of basic open sets U_j^*, on each of which the oscillation of $f(y)$ is less than epsilon. The corresponding sets U_j of X are Z-complements, they cover X, and on each of them the oscillation of the restriction $f(x)$ of $f(y)$ to X is less than epsilon. Hence every such function $f(x)$ defined over X and having a continuous extension to $w(Z)$ is Z-uniformly continuous.

Conversely, suppose the real function $f(x)$ is Z-uniformly continuous on a semi-normal space X with normal base Z. We may assume without loss of generality that Z contains all singletons $\{x\}$ of X, since if Z is extended to the smallest ring of sets containing Z and all singletons, a new normal base will be obtained leading to the same compactification, and with the same determination of Z-uniform continuity.

We define a function $g(y)$ over $w(Z)$ which is intended to be a continuous extension of $f(x)$, as follows. Each point y of $w(Z)$ is a Z-ultrafilter, and therefore a family of sets of Z which is dually directed by the subset relation. From every member n of the family of sets y we select a point x_n of X which is a member of n. This is possible by the axiom of choice, since no member of y is empty. Consider the net $\{f(x_n)\}$, with indices n ordered by the superset relation. Since $f(x)$ is Z-uniformly continuous, it follows that the net $\{f(x_n)\}$ satisfies the Cauchy convergence criterion for nets of real numbers. Here this net converges to a real number which we call $g(y)$. If y is the image under the imbedding map of a point x of X, then y as a directed set has the last element $\{x\}$, and the net converges to $f(x)=g(y)$. Hence g is an extension of f.

We now show that $g(y)$ is continuous. Given $\epsilon > 0$, let the family $\{U_j\}$ be a finite open cover of X by Z-complements, on each member of which the oscillation of $f(x)$ is less than $\epsilon/3$. The corresponding family $\{U_j^*\}$ is an open cover of $w(Z)$. It follows that the oscillation of $g(y)$ on each set U_j^* of the cover is less than ϵ. For let y and y' be any two Z-ultrafilters of U_j^*. Then by definition of $g(y)$ there exist elements m and n of y and y' respectively, each a subset of U_j, such that $|f(x_m)-g(y)|<\epsilon/3$ and $|f(x_n)-g(y')|<\epsilon/3$. Hence $|g(y)-g(y')|<\epsilon$, and the function $g(y)$ is continuous. This completes the proof.

Theorem 2 can be generalized to give a necessary and sufficient condition that a continuous map of a semi-normal space X with normal base Z into a compact Hausdorff space H can be extended to a continuous map of the compactification $w(Z)$ into H. We say that a map of X into H is *uniformly continuous relative to Z and H* if corresponding to every finite open cover F of H there exists a finite open cover G of X by sets which are Z-complements, such that each member of G is sent into a subset of a member of F by the map.

A proof similar to that of Theorem 2 will then show that a map of X into H can be extended to a continuous map of $w(Z)$ into H if and only if the map is uniformly continuous relative to Z and H.

Theorem 2 suggests that the concept of a normal base Z for a Tychonoff space X is similar to that of a uniformity for X. It suggests also a modification of the Čech homology theory for a space with normal base Z, in which the finite open covers are restricted to those whose members are Z-complements.

Conclusion. It is clear that different compactifications $w(Z)$ of Wallman type may be obtained for a non-compact Tychonoff space X by choosing different normal bases Z for X. The question remains open whether every Hausdorff compactification Y of X can be obtained in this way, in the sense that Y is homeomorphic to some Wallman space $w(Z)$. Every compactification Y of X determines a normal base for X consisting of the zero-sets of the ring of those continuous real functions defined over X which may be extended to continuous functions over Y. However, examples show that the Wallman space $w(Z)$ determined by the zero-sets of functions continuously extendable to Y may not be homeomorphic to Y.

Consider the one-point compactification N^* of the countable discrete space N. It may easily be seen that every subset of N is the zero-set of some real function continuously extendable to N^*. Hence the Wallman space determined by these zero-sets is the Stone-Čech space βN, which is certainly not homeomorphic to N^*. This example is due to Robert M. Brooks of the University of Minnesota.

However, Brooks has shown that the one-point Hausdorff compactification of a locally compact Hausdorff space X may always be obtained as a

Wallman space $w(Z)$, where the normal base Z consists of the zero-sets of those continuous real functions on X which are constant on some complement of a compact set. The results of Brooks have not yet been published.

Every Hausdorff compactification Y of a semi-normal space X may be represented as the structure space of maximal ideals of the ring R of those real functions in X which are continuously extendable to Y; it may also be represented as a structure space of those subrings of R which are lattice ordered and which distinguish points of Y. It is conceivable, but by no means evident, that Y may be always isomorphic to the Wallman space $w(Z)$ determined by the normal base Z consisting of the zero-sets of some subring of R.

It should be noted that a normal base Z for a semi-normal space does not necessarily consist of zero-sets. Given a normal base Z of zero-sets of a space X, it often happens that a subfamily of Z of lower cardinality than Z exists which leads to the same Wallman compactification as Z.

Normality is neither hereditary nor productive; subspaces and cartesian products of normal spaces need not be normal. However, the property of being semi-normal is both hereditary and productive since it is equivalent to being T_1 and completely regular, and these properties are hereditary and productive. It would be desirable to have simple direct proofs that subspaces and cartesian products of spaces with a normal base have normal bases.

THE PENNSYLVANIA STATE UNIVERSITY,
UNIVERSITY PARK, PA.

REFERENCES.

[1] P. S. Aleksandrov, "On completely regular spaces and their bicompact extensions," *Vestnik Moskovskogo Universiteten Ser.* 1, No. 2, (1962) pp. 37-43.
[2] Orrin Frink, "Ideals in partially ordered sets," *American Mathematical Monthly,* vol. 61 (1954), pp. 223-234.
[3] L. Gillman and M. Jerison, *Rings of Continuous Functions,* Princeton, 1960.
[4] J. L. Kelley, *General Topology,* Princeton, 1955.
[5] W. J. Pervin, "On separation and proximity spaces," *American Mathematical Monthly,* vol. 71 (1964), pp. 158-161.
[6] Pierre Samuel, "Ultrafilters and compactification of uniform spaces," *Transactions of the American Mathematical Society,* vol. 64 (1948), pp. 100-132.
[7] M. H. Stone, "Applications of the theory of Boolean rings to general topology," *Transactions of the American Mathematical Society,* vol. 41 (1937), pp. 375-481.
[8] A. D. Wallace, "Separation spaces II," *Annais da Academia Brasileira de Ciencas,* vol. 14 (1942), pp. 203-206.
[9] Henry Wallman, "Lattices and topological spaces," *Annals of Mathematics* (2), vol. 42 (1938), pp. 687-697.

FINITE GROUPS HAVING A FAITHFUL REPRESENTATION OF DEGREE LESS THAN $(2p+1)/3$.

By DAVID L. WINTER.[1]

1. Introduction.

The principal result of this paper is the

MAIN THEOREM. *Let p be a prime and let G be a finite group of order $p^a g'$ where $(p, g') = 1$. If G has a faithful representation $\mathfrak{z}$ of degree n over the complex number field, and if $n < (2p+1)/3$, then one of the following holds:*

(A) *The p-Sylow subgroup of G is an abelian normal subgroup of G.*

(B) $p > 7$. *$G = HM$ is the product of normal subgroups H and M of G where* (i) $G/H \cong LF(2, p)$. (ii) *H contains an abelian normal p-Sylow subgroup P_0 of order p^{a-1}.* (iii) *$M/B \cong LF(2,p)$ where B is the normal subgroup of M consisting of all p-regular elements of M which are commutative with an arbitrary fixed element of M of order p.* (iv) $H \cap M = B$. (v) *The centralizer of P_0 in G is $P_0 \times M$.* (vi) *If P_1 is an arbitrary p-Sylow subgroup of M, then the centralizer of P_1 in G is $P_1 \times H$. Further, if $\mathfrak{z}$ is irreducible or if $n = (p \pm 1)/2$, then $H = Z$, the center of G.*

(C) $p = 7$. $G/Z \cong LF(2, 7)$ *or* A_7.

(D) $p = 5$. $G/Z \cong LF(2, 5)$ *or* A_6.

(E) $p = 3$. $G/Z \cong LF(2,3)$, S_4 *or* A_5.

A first step here will be to prove

THEOREM I. *Let p be a prime and let G be a finite group which has a faithful representation of degree $n = (p-1)/2$ over the complex number field. Then either G has an abelian normal p-Sylow subgroup or the factor group, G/Z, of G modulo the center of G is isomorphic to $LF(2, p)$.*

These theorems have already been proved in the case when the order of G is divisible by the prime p to the first power only: Theorem I by R. Brauer

---

Received January 7, 1964.

[1] Presented to the American Mathematical Society in April, 1963. A part of this paper was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Carnegie Institute of Technology.

([2], Theorem 4) and the Main Theorem by H. F. Tuan [8]. If in the statement of Theorem I $n = (p-1)/2$ is replaced by $n < (p-1)/2$, W. Feit and J. G. Thompson [5] have shown that the group G has an abelian normal p-Sylow subgroup. This result as well as the methods of [5] are used in the proof of Theorem 1. The proof of the Main Theorem is made to depend on a special case of the following result due to W. Feit [4]:

(1A) *Let G be a group of order g which has a faithful representation of degree n over the complex numbers. Let $g = g_1h$ where $(g_1, h) = 1$ and $p > n+1$ for every prime p dividing h. Then G contains a normal abelian subgroup K of order h or h/p for some prime p.*

The author would like to acknowledge the guidance of his thesis advisor, Professor H. S. Leonard, Jr., during this research work.

2. Notation and preliminary results. All groups considered in this paper are of finite order and all group representations are representations by linear transformations over the complex number field.

$H \lhd G$ means that H is a normal subgroup of G. $Z(H)$ is the center of the subgroup H and we shall write $Z = Z(G)$. $|G|$ denotes the order of G. If S is an arbitrary subset of G, $C(S)$ and $N(S)$ denote, respectively, the centralizer of S in G and the normalizer of S in G. $C(y)$ is the centralizer of the element y in G. If S and T are subsets of the group G, $S - T$ denotes the set of all elements in S which are not members of T.

The following short results are used frequently in later sections.

(2.1) *Let p be a prime. If the finite group G has a faithful representation of degree $n < (p-1)$ over the complex number field and if G contains a normal p-complement, then G has a normal p-Sylow subgroup.*

Proof. Let $|G| = p^a g'$ with $(p, g') = 1$. By assumption G contains a normal subgroup H of order $|H| = g'$. Let q be a prime divisor of g' and let Q be a Sylow q-subgroup of H. Every conjugate of Q relative to G is a conjugate of Q relative to H; hence, $(G\colon N(Q)) = (H\colon H \cap N(Q))$. The latter is relatively prime to p and therefore $p^a \mid |N(Q)|$. Thus if P is a Sylow p-subgroup of G, Q may be chosen so that $P \subseteq N(Q)$. Then PQ is a group with $Q \lhd PQ$ and PQ is solvable. N. Ito [6] has proved that a finite solvable group having a faithful representation of degree $n < (p-1)$ over the complex number field contains a normal p-Sylow subgroup. Therefore $P \lhd PQ$ and $Q \subseteq N(P)$. Since q was an arbitrary prime dividing g', we see that $g' \mid |N(P)|$ which implies $N(P) = G$.

(2.2) *Let p be a prime and let G be a finite group which has a faithful representation $\mathfrak{z}$ of degree n over the complex number field. If $n < p$, then every p-Sylow subgroup of G is abelian.*

For, let P be a p-Sylow subgroup of G. Then $\mathfrak{z} \mid P$ is a faithful representation of P. Since $n < p$ and since the degree of each irreducible constituent of $\mathfrak{z} \mid P$ is a factor of $|P|$, each of these degrees must be 1. Therefore $\mathfrak{z} \mid P$, considered as a matrix representation of P, in completely reduced form is an isomorphism of P onto a group of diagonal matrices. Hence P is abelian.

The result of Feit and Thompson [5] is easily generalized in the following way.

(2.3) *Let p be a prime and let G be a finite group which has a faithful representation $\mathfrak{z}$ over the complex number field all of whose irreducible constituents are of degree less than $(p-1)/2$. Then G has an abelian normal p-Sylow subgroup.*

Proof. Let $\mathfrak{D}_1, \mathfrak{D}_2, \cdots, \mathfrak{D}_t$ be the irreducible constituents of $\mathfrak{z}$ and consider $\mathfrak{z}$ to be a matrix representation of G. Then for all $x \in G$ the matrices $[\mathfrak{z}(x)]$, in completely reduced form appear as

$$[\mathfrak{z}(x)] = \begin{bmatrix} [\mathfrak{D}_1(x)] & & & 0 \\ & [\mathfrak{D}_2(x)] & & \\ & & \ddots & \\ 0 & & & [\mathfrak{D}_t(x)] \end{bmatrix}$$

where the 0's indicate that the remaining entries are all zeros. Let $K_1, K_2, \cdots, K_t$ be the kernels of the representations $\mathfrak{D}_1, \mathfrak{D}_2, \cdots, \mathfrak{D}_t$, respectively. Then $\mathfrak{D}_i$ is a faithful representation of the factor group G/K_i of degree less than $(p-1)/2$ for $i = 1, 2, \cdots, t$. By the result of [5] G/K_i has an abelian normal p-Sylow subgroup for all i. Now let x and y be any two elements of G of p-power order. Then the cosets K_ix and K_iy are elements of G/K_i of p-power order. Hence the matrices $[\mathfrak{D}_i(x)]$ and $[\mathfrak{D}_i(y)]$ commute for $i = 1, 2, \cdots, t$. It follows by block multiplication of matrices that $[\mathfrak{z}(x)]$ and $[\mathfrak{z}(y)]$ commute. Since x and y were arbitrary elements of G of p-power order, it follows that the set of all elements of G of p-power forms an abelian group, hence an abelian normal p-Sylow subgroup of G.

(2.4) *Let G be a finite group which has a faithful representation $\mathfrak{z}$ over the complex number field and let H be a subgroup of G. If $\mathfrak{z} \mid H$ is an irreducible representation of H, then $Z(H) \subseteq Z$.*

For, $\mathfrak{z}$ may be considered to be a representation of G by matrices. If $\mathfrak{z} \mid H$ is irreducible, then from Schur's lemma each member of $Z(H)$ is represented by a scalar multiple of the identity matrix. Hence, the matrices corresponding to $Z(H)$ are commutative with all the matrices corresponding to elements of G. Since $\mathfrak{z}$ is one-one, this implies that $Z(H) \subset Z$.

The following result is an immediate consequence of Theorem 7, page 173, of [9].

(2.5) *Let p be a prime and let G be a finite group which has an abelian p-Sylow subgroup. If $Z = Z(G)$ and $p \mid |Z|$, then G contains a proper normal subgroup whose index in G is a power of p.*

(2.6) *Let p be a prime and let P be a p-Sylow subgroup of the finite group G. If K is a subgroup of G with $K \subseteq Z$ and if G/K has a normal p-Sylow subgroup, then $P \triangleleft G$.*

For, if G/K has a normal p-Sylow subgroup, then KP/K is this Sylow subgroup and $KP \triangleleft G$. Since $K \subseteq Z$, $P \triangleleft KP$. Hence, $P \triangleleft G$.

3. Proof of Theorem I. The following result was proved in [5] although it was not explicitly stated there.

(3A) *Let p be a prime. If the finite group G has a faithful representation of degree $n < (p-1)$ over the complex number field, then p^2 does not divide the order of G if the following conditions hold:* (i) *The center Z of G is the unique maximal normal subgroup of G and $p \nmid |Z|$;* (ii) *Let P be a p-Sylow subgroup of G. Let N be the set of all elements in the normalizer $N(P)$ of P which are not in the center of G but whose centralizers do contain elements of P other than the identity. Then for $x \in G$, $N \cap xNx^{-1}$ not empty implies $x \in N(P)$.*

Throughout the remainder of this section it will be assumed that G is a counter-example to Theorem I of minimal order. This implies, in particular, that Theorem I is true for every proper subgroup of G and that $p > 3$. By means of a series of steps similar to those in [5] it will be shown that G satisfies all the hypotheses of (3A) and, hence, that $p^2 \nmid |G|$. Since Theorem I has been proved in the case when the order of G is divisible by p to the first power only, this will yield a contradiction which suffices to complete the proof.

P will always denote a fixed p-Sylow subgroup of G. By (2.2)

(3.1) *P is an abelian group.*

(3.2) *The index of a proper normal subgroup of G cannot be a power of p.*

Let H be a counter-example to (3.2). Assume first that H has a normal p-Sylow group P_H (the case $P_H = \{1\}$ is included here). Since $H \lhd G$, $P_H \lhd G$ and therefore $C(P_H) \lhd G$. Since P is abelian, $P \subseteq C(P_H)$. Hence, $C(P_H)$ cannot have a normal p-Sylow subgroup, since the same would then be true of G. Since the faithful representation $\mathfrak{z}$ of G has degree $(p-1)/2$, it follows from (2.3) that $\mathfrak{z} \mid C(P_H)$ is an irreducible representation of $C(P_H)$. By (2.4) $P_H \subseteq Z$. By a theorem due to W. Burnside H contains a normal p-complement which is necessarily a normal p-complement of G. (2.1) now yields that $P \lhd G$, contrary to assumption. Therefore, H does not have a normal p-Sylow subgroup.

The minimal nature of G now yields that $H/Z(H) \cong LF(2,p)$. By (2.3) $\mathfrak{z} \mid H$ must be irreducible and hence by (2.4), $Z(H) \subseteq Z$. $H/Z(H)$ is thus a normal subgroup of $G/Z(H)$ whose index in $G/Z(H)$ is a positive power of p. M. Suzuki ([7], Lemma 2) has shown that the index of the group of inner automorphisms of $LF(2,p)$ in the group of all automorphisms of $LF(2,p)$ is 2. It follows easily from this result that, if L is a subgroup of an arbitrary finite group with $L \cong LF(2,p)$, then the index of $C(L)L$ in $N(L)$ is 1 or 2. Since $(G : H)$ is a positive power of p, it may be concluded that if R is the normal subgroup of G, containing $Z(H)$, such that $R/Z(H)$ is the centralizer of $H/Z(H)$ in $G/Z(H)$, then $(R/Z(H))(H/Z(H)) = G/Z(H)$. Since $H/Z(H) \cong LF(2,p)$ is a simple group, $R/Z(H) \cap H/Z(H) = \{1\}$ and hence $|R/Z(H)| = (G/Z(H) : H/Z(H))$ which is a positive power of p. Therefore $|R| = p^k |Z(H)|$ for some $k > 0$. Since $Z(H) \subseteq Z$, if P_R is a p-Sylow subgroup of R, $R = P_R Z(H)$. But $Z(H) \subseteq N(P_R)$ so that $P_R \lhd P_R Z(H) = R$. Since $R \lhd G$, $P_R \lhd G$ and $C(P_R) \lhd G$. P being abelian implies that $P \subseteq C(P_R)$ and, thus, $C(P_R)$ cannot have a normal p-Sylow subgroup or the same would be true of G. (2.3) yields that $\mathfrak{z} \mid C(P_R)$ is irreducible and (2.4) that $P_R \subseteq Z$. Therefore $R = P_R Z(H) \subseteq Z$. But $G/R = RH/R \cong H/R \cap H = H/Z(H) \cong LF(2,p)$ is a simple group. Hence, R is a maximal normal subgroup of G. Therefore, $R = Z$ and $G/Z \cong LF(2,p)$, a contradiction. This completes the proof of (3.2).

(3.3) *Z is the unique maximal normal subgroup of G. $p \nmid |Z|$.*

Let H be a proper normal subgroup of G. Then PH is a group and by (3.2) $PH \neq G$. Suppose that PH does not have a normal p-Sylow subgroup. Then by the induction hypothesis $PH/Z(PH) \cong LF(2,p)$. Since the order of $LF(2,p)$ is divisible by p to the first power only, $|Z(PH)|$ is divisible by

p^{a-1} where p^a is the highest power of p dividing $|G|$. But $a \geqq 2$ since Theorem I has been proved in the case that $|G|$ is divisible by p to the first power only. Therefore, $p^{a-1} \geqq p$. (2.3) and (2.4) yield as before that $Z(PH) \subseteq Z$ and hence $p \mid |Z|$. (2.5) now leads to a contradiction of (3.2). Therefore, PH and, consequently, H contain normal p-Sylow subgroups. If P_H is the p-Sylow subgroup of H, then $P_H \lhd G$. Since P is abelian, $P \subseteq C(P_H) \lhd G$. P is not a normal subgroup of $C(P_H)$ for otherwise it would follow that $P \lhd G$. Application of (2.3) and (2.4) yield that $P_H \subseteq Z$. Hence, $P_H = \{1\}$ for if not (2.5) leads to a contradiction of (3.2). It has thus been shown that, if H is a proper normal subgroup of G, then $p \nmid |H|$ and that $P \lhd PH$.

Since $P \cap H = \{1\}$, $PH = P \times H$ and therefore, $P \subseteq C(H)$. Since $C(H) \lhd G$, we must have $C(H) = G$. Therefore, $H \subseteq Z$ and (3.3) follows.

(3.4) $P \cap xPx^{-1} = \{1\}$, *unless* $x \in N(P)$.

For, let $P_0 = P \cap xPx^{-1}$ be an intersection of distinct p-Sylow subgroups of G. Then both P and xPx^{-1} are contained in $C(P_0)$ so that $C(P_0)$ does not have a normal p-Sylow subgroup. By (2.3) $\mathfrak{z} \mid C(P\ )$ is irreducible. Hence, (2.4) yields that $P_0 \subseteq Z$ and (3.3) that $P_0 = \{1\}$.

Let N be the set of all elements of $N(P) - Z$ whose centralizers contain elements of P other than the identity.

(3.5) $N \cap xNx^{-1} \neq \phi$ *implies* $x \in N(P)$.

For, if $y \in N \cap xNx^{-1}$, then both y and $x^{-1}yx$ are members of N. By the definition of N there exist $P_1 \neq \{1\}$ and $P_2 \neq \{1\}$, subgroups of P, such that $P_1 \subseteq C(y)$ and $P_2 \subseteq C(x^{-1}yx)$, or such that both P_1 and xP_2x^{-1} are contained in $C(y)$. Since $y \notin Z$, (2.4) implies that $\mathfrak{z} \mid C(y)$ is reducible and (2.3) that $C(y)$ has a normal p-Sylow subgroup. Since P_1 and xP_2x^{-1} are both contained in this Sylow subgroup, (3.4) implies that they are both contained in P. Hence by (3.4) $x \in N(P)$.

(3.3) and (3.5) show that G satisfies the hypotheses of (3A). As mentioned above this is sufficient to complete the proof of Theorem I.

4. Proof of the Main Theorem when $p > 7$. The following assumptions are made throughout the next four lemmas.

p is a prime with $p > 7$. G is a finite group of order $p^a g'$ with $(p, g') = 1$ which has a faithful representation $\mathfrak{z}$ of degree $n < (2p+1)/3$ over the complex number field. G contains a normal subgroup P_0 of order p^{a-1} but G does not contain a normal p-Sylow subgroup.

(4.1) *If P_0 is contained in the center Z of G, then G is of the form $G = P_0 \times M$, where $M/B \cong LF(2,p)$, B being the normal subgroup of M consisting of all p-regular elements of M which are commutative with an arbitrary fixed element of M or order p.*

Proof. Let G be a counter-example to (4.1) of minimal order. Then necessarily $a \geqq 2$, for if $p^2 \nmid |G|$, Tuan's Main Theorem yields a contradiction. Consequently, $|P_0| = p^{a-1} \geqq p$ and by (2.5) G contains a proper normal subgroup H whose index in G is a power of p. Suppose H has a normal p-Sylow subgroup P_H. Since $H \lhd G$, $P_H \lhd G$ and $C(P_H) \lhd G$. By (2.2) every p-Sylow subgroup of G is abelian so that $p^a \mid |C(P_H)|$. Therefore, if $C(P_H)$ had a normal p-Sylow subgroup, the same would be true of G. Thus, $C(P_H)$ does not have a normal p-Sylow subgroup. Let $H_1 = C(P_H) \cap H$. H_1 contains all p-regular elements of $C(P_H)$ since H contains all p-regular elements of G. P_H is a p-Sylow subgroup of H_1 contained in the center of H_1. By Burnside's theorem H_1 contains a normal p-complement which is necessarily a normal p-complement of $C(P_H)$. Therefore, by (2.1) $C(P_H)$ has a normal p-Sylow subgroup which is a contradiction. Hence, H does not have a normal p-Sylow subgroup. Let $|H| = p^b g'$. Then since $G = P_0 H$, $|P_0 \cap H| = |P_0| \cdot |H| / |G| = p^{a-1} p^b g' / p^a g' = p^{b-1}$. Therefore, $P_1 = P_0 \cap H$ is a p-Subgroup of H of order p^{b-1} contained in the center of H. The minimal nature of G yields that $H = P_1 \times M$ where M is as described in (4.1). Since $P_0 \cap M = \{1\}$, $G = P_0 \times M$.

(4.2) *If $\mathfrak{z}$ is irreducible, then $G/Z \cong LF(2,p)$ and $P_0 \subseteq Z$.*

Proof. Since $P_0 \lhd G$, $C(P_0) \lhd G$ and since every p-Sylow subgroup of G is abelian, $p^a \mid |C(P_0)|$. Hence, $C(P_0)$ cannot have a normal p-Sylow subgroup or else the same would be true of G. Suppose that $\mathfrak{z} \mid C(P_0)$ is reducible. Then, by a theorem due to A. H. Clifford [3], all irreducible constituents of $\mathfrak{z} \mid C(P_0)$ must have the same degree. Hence, all irreducible constituents of $\mathfrak{z} \mid C(P_0)$ have degree less than $(p-1)/2$. (2.3) now yields that $C(P_0)$ has a normal p-Sylow subgroup which is a contradiction. Therefore, $\mathfrak{z} \mid C(P_0)$ is an irreducible representation of $C(P_0)$. It follows from (2.4) that $P_0 \subseteq Z$. From (4.1) $G = P_0 \times M$, where $M/B \cong LF(2,p)$. Now $\mathfrak{z} \mid M$ is irreducible because of the theorem about irreducible representations of direct products. Consequently, Tuan's Main Theorem yields that $B = Z(M)$, the center of M. Since $G = P_0 \times M$, $Z = P_0 Z(M)$, and

$$G/Z = P_0 M / P_0 Z(M) \cong M/Z(M) \cong LF(2,p).$$

(4.3) *If the degree n of $\mathfrak{z}$ is $(p \pm 1)/2$, then $G/Z \cong LF(2,p)$.*

Proof. If $n=(p-1)/2$, the result follows from Theorem I and if $\mathfrak{z}$ is irreducible, (4.3) follows from (4.2). It may therefore be assumed that $\mathfrak{z}$ has degree $n=(p+1)/2$ and is reducible. Since G does not have a normal p-Sylow subgroup, we conclude from (2.3) that $\mathfrak{z}$ is the sum of two irreducible constituents, $\mathfrak{D}$ of degree 1 and $\mathcal{E}$ of degree $(p-1)/2$. Thus, in completely reduced matrix form $\mathfrak{z}$ appears as

$$[\mathfrak{z}(x)]=\begin{bmatrix} [\mathfrak{D}(x)] & 0\cdots 0 \\ 0 & \\ \vdots & [\mathcal{E}(x)] \\ 0 & \end{bmatrix}.$$

It is clear, using Schur's lemma, that the center Z of G is the set of all x such that $[\mathcal{E}(x)]$ is a scalar matrix. Hence if K is the kernel of $\mathcal{E}$, then $\mathcal{E}$ is a faithful representation of G/K and $K\subseteq Z$. G/K does not have a normal p-Sylow subgroup because otherwise G would have a normal p-Sylow subgroup by (2.6). It follows from Theorem I that if L/K is the center of G/K, then $(G/K)/(L/K)\cong G/L\cong LF(2,p)$. But L is the set of all $x\in G$ for which $[\mathcal{E}(x)]$ is a scalar matrix. Hence $L=Z$ and $G/Z\cong LF(2,p)$.

(4.4) $G=HM$ *with* $H\lhd G$, $M\lhd G$ *where* (i) $G/H\cong LF(2,p)$. (ii) P_0 *is the normal p-Sylow subgroup of* H. (iii) $M/B\cong LF(2,p)$ *where* B *is the set of all p-regular elements of* M *which are commutative with an arbitrary fixed element of* M *of order* p. (iv) $H\cap M=B\lhd G$. (v) $C(P_0)=P_0\times M$. (vi) *If* P_1 *is an arbitrary p-Sylow subgroup of* M, *then* $C(P_1)=P_1\times H$.

Proof. Suppose that $\mathfrak{z}$ is irreducible. Then by (4.2) and (4.1) $G/Z\cong LF(2,p)$, $P_0\subseteq Z$ and $G=P_0\times M$. Here $M/B\cong LF(2,p)$ where B is defined in the statement (4.1). But since by the theorem on irreducible representations of direct products $\mathfrak{z}\mid M$ is irreducible, Tuan's Main Theorem yields that $B=Z(M)$. Since $G=P_0\times M$, $Z=P_0Z(M)$. Set $H=P_0Z(M)$. Then $G=HM$ and all conclusions of (4.4) follow. Hence it may be assumed that $\mathfrak{z}$ is reducible.

Since G does not have a normal p-Sylow subgroup, it follows from (2.3) that $\mathfrak{z}$ has an irreducible constituent, say $\mathfrak{z}_1$, of degree $\geqq(p-1)/2$. Let $\mathfrak{z}_2$ be the sum of the remaining constituents so that $\mathfrak{z}=\mathfrak{z}_1\oplus\mathfrak{z}_2$. Consider the kernels K_1 and K_2 of $\mathfrak{z}_1$ and $\mathfrak{z}_2$. $\mathfrak{z}_1$ and $\mathfrak{z}_2$ are faithful representations of G/K_1 and G/K_2, respectively. Since $n<(2p+1)/3$, the degree of each irreducible constituent of $\mathfrak{z}_2$ must be $<(p-1)/2$ so that by (2.3) G/K_2 has a normal abelian p-Sylow subgroup. Further, G/K_1 does not have a normal p-Sylow subgroup for otherwise an argument like that in the proof of (2.3) would lead to the contradiction that G has a normal p-Sylow subgroup. Since $\mathfrak{z}$ is faithful, $K_1 \cap K_2 = \{1\}$ and hence $\mathfrak{z}_2 \mid K_1$ is a faithful representation of K_1 all of whose irreducible constituents are of degree $< (p-1)/2$. By (2.3) K_1 has a normal p-Sylow subgroup P_1. Since $K_1 \lhd G$, $P_1 \lhd G$. P_1 is therefore contained in the intersection Q of all p-Sylow subgroups of G. Since $Q \lhd G$, $|Q| < p^a$. But $P_0 \subseteq Q$ and has order p^{a-1}. Hence, $Q = P_0$ and $P_1 \subset P_0$. Let P be a p-Sylow subgroup of G. Then P_0 and P_1 are both contained in P and $P_1 = P \cap K_1 = P_0 \cap K_1$. Now PK_1/K_1 is a p-Sylow subgroup of G/K_1 and is not normal in G/K_1, but P_0K_1/K_1 is a normal subgroup of G/K_1. Let the order of PK_1/K_1 be p^{a-b}. Then $|P_0K_1/K_1| = |P_0/K_1 \cap P_0| = |P_0/K_1 \cap P| = p^{(a-b)-1}$. Since $\mathfrak{z}_1$ is irreducible, this shows that the result of (4.2) may be applied to the group G/K_1. We conclude that if H is the normal subgroup of G such that H/K_1 is the center of G/K_1, then $G/H \cong (G/K_1)/(H/K_1) \cong LF(2,p)$.

Since the order of $LF(2,p)$ is divisible by p to the first power only, a p-Sylow subgroup of H must have order p^{a-1}. Because of the fact that $P_0H \lhd G$ and H is a maximal normal subgroup of G, $P_0H = H$. Hence $P_0 \subseteq H$ and P_0 is the normal p-Sylow subgroup of H.

Since $C(P_0)$ does not have a normal p-Sylow subgroup, (4.1) yields that $C(P_0) = P_0 \times M$ where $M/B \cong LF(2,p)$. $C(P_0) \lhd G$ implies that for any $x \in G$, $xMx^{-1} \subseteq C(P_0)$ and $xMx^{-1} \lhd C(P_0)$ since $M \lhd C(P_0)$. $M \cap xMx^{-1} \lhd C(P_0)$ and contains all p-regular elements of $C(P_0)$. Hence B is properly contained in $M \cap xMx^{-1}$ and since B is a maximal normal subgroup of M, we must have $M \cap xMx^{-1} = M$. Thus $M = xMx^{-1}$ and $M \lhd G$. The definition of B implies that B is the normal subgroup of M consisting of all p-regular elements of M which are commutative with all elements of M of order p. Hence B is a characteristic subgroup of M, so that $B \lhd G$. Recalling $H \lhd G$, we have $HB \lhd G$ which implies that $B \subseteq H$, H being a maximal normal subgroup of G. Therefore $B \subseteq H \cap M \lhd M$. $H \cap M = M$ is impossible since H has a normal p-Sylow subgroup and M does not. Since B is a maximal normal subgroup of M, we conclude that $H \cap M = B$. Hence, $HM/H \cong M/H \cap M = M/B \cong LF(2,p)$ and since $G/H \cong LF(2,p)$, we conclude that $G = HM$. Further, $(H/B)(M/B) = (H/B) \times (M/B)$ and therefore H/B centralizes M/B.

From these remarks it follows easily that if P_1 is an arbitrary p-Sylow subgroup of M, then $P_1H = P_1 \times H \subseteq C(P_1)$. On the other hand, an element of $C(P_1)$ has the form hm with $h \in H$, $m \in M$ since $G = HM$. But for such an element $m \in C(P_1) \cap M = P_1 \times B \subseteq P_1 \times H$. Therefore $hm \in P_1 \times H$,

$C(P_1) \subseteq P_1 \times H$ and hence $C(P_1) = P_1 \times H$. This completes the proof of (4.4).

If in (4.4) $\mathfrak{z}$ is irreducible or has degree $(p \pm 1)/2$, it follows from (4.3) and (4.2) that $G/Z \cong LF(2,p) \cong G/H$. In these cases Z and H are both maximal normal subgroups of G and hence, since $HZ \neq G$, $H = Z$. Therefore the Main Theorem has been established for the case $p > 7$ if it is known that a finite group of order $p^a g'$ with $(p, g') = 1$ and $p > 7$ which has a faithful representation of degree $n < (2p + 1)/3$ over the complex number field has a normal abelian subgroup of order p^a or p^{a-1}. This fact readily follows from (1A).

5. **Proof of the Main Theorem when $p \leqq 7$.** If p is a prime with $p \leqq 7$, then the group G under consideration is a finite linear group over the complex number field of degree $n \leqq 4$. In his book [1] H. F. Blichfeldt has listed the collineation groups corresponding to all finite linear groups of degrees $n = 2$, 3 and 4 over the complex number field. His results will be applied after some reductions of the problem.

Now let G satisfy the hypotheses of the Main Theorem. Since $n < (2p + 1)/3 < p$, (2.2) implies that a Sylow p-subgroup of G is abelian. If $p = 2$, then n must be 1 and the result follows since G is abelian.

If $p = 3$, then $n < 7/3 < 3$, so that only the case $n = 2$ need be considered. Clearly, $\mathfrak{z}$ may be assumed to be irreducible for otherwise G would have a faithful representation by 2×2 diagonal matrices and would be abelian. Hence, Schur's lemma yields that the center Z of G consists of all x for which $[\mathfrak{z}(x)]$ is a scalar matrix. Thus, G/Z coincides with the collineation group of G. But from Blichfeldt's list it is clear that either G/Z has a normal 3-Sylow subgroup or is isomorphic to one of $LF(2,3)$, S_4 or A_5. By (2.6) this yields the desired result for G.

If $p = 5$, then $n < 11/3 < 4$. If $n = 2 = (5 - 1)/2$ the result follows from Theorem I. Hence it may be assumed that the faithful representation $\mathfrak{z}$ of G has degree 3. Suppose first that $\mathfrak{z}$ is reducible. If $\mathfrak{z}$ has three irreducible constituents, then G would have a faithful representation by 3×3 diagonal matrices and, hence, would be abelian. Therefore, assume that $\mathfrak{z}$ is the direct sum of two irreducible constituents $\mathfrak{z}_1$ and $\mathfrak{z}_2$, of degrees 1 and 2, respectively. Let K be the kernel of $\mathfrak{z}_2$. Then $\mathfrak{z}_2$ is a faithful representation of G/K. Let L be the normal subgroup of G such that L/K is the center of G/K. Then L is the set of all $x \in G$ such that $[\mathfrak{z}_2(x)]$ is a scalar matrix and therefore $L = Z$. By Theorem I either G/K has a normal 5-Sylow subgroup or $G/L \cong (G/K)/(L/K) \cong LF(2,5)$. Since $K \subseteq L = Z$, the desired result follows from (2.6). It may therefore be assumed that $\mathfrak{z}$ is an irreducible representation of G.

G/K then coincides with the collineation group of G. Blichfeldt's list shows that this collineation group either has an abelian normal Sylow 5-subgroup or $G/Z \cong LF(2,5)$ or A_6. By (2.6) this yields the desired result.

When $p=7$, $n<(2p+1)/3=5$. If $n=2<3=(p-1)/2$, then G has an abelian normal 7-Sylow subgroup by (2.3). If $n=3$, the Main Theorem follows from Theorem I. Hence suppose that G has a faithful representation $\mathfrak{z}$ of degree 4. If $\mathfrak{z}$ is reducible and if every irreducible constituent of $\mathfrak{z}$ has degree less than $3=(p-1)/2$, then (2.3) yields that G has an abelian normal 7-Sylow subgroup. Therefore it may be assumed that if $\mathfrak{z}$ is reducible, then it is the direct sum of two irreducible constituents, $\mathfrak{z}_1$ and $\mathfrak{z}_2$, of degrees 1 and 3, respectively. Let K be the kernel of $\mathfrak{z}_2$. Then by Theorem I G/K either has an abelian normal 7-Sylow subgroup or the factor group $(G/K)/(L/K)$ of G/K modulo the center of G/K is isomorphic to $LF(2,7)$. The treatment of these two cases is similar to that used in the case $p=5$.

It may therefore be assumed that $\mathfrak{z}$ is a faithful irreducible representation of degree 4. Hence G/Z is the collineation group of G. From Blichfeldt's results this collineation group either has an abelian normal 7-Sylow subgroup or is isomorphic to $LF(2.7)$ or A_7. By (2.6) this suffices to give the desired result. The proof of the Main Theorem is now complete in all cases.

CARNEGIE INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] H. F. Blichfeldt, *Finite Collineation Groups*, Chicago, 1917.

[2] R. Brauer, "On groups whose orders contain a prime number to the first power II," *American Journal of Mathematics*, vol. 64 (1942), pp. 421-440.

[3] A. H. Clifford, "Representations induced in an invariant subgroup," *Annals of Mathematics*, vol. 38 (1937), pp. 533-550.

[4] W. Feit, "Groups which have a faithful representation of degree less than $p-1$," *to appear*.

[5] W. Feit and J. G. Thompson, "Groups which have a faithful representation of degree less than $(p-1)/2$," *Pacific Journal of Mathematics*, vol. 11 (1961), pp. 1257-1262.

[6] N. Ito, "On a theorem of H. F. Blichfeldt," *Nagoya Mathematical Journal*, vol. 5 (1954), pp. 75-77.

[7] M. Suzuki, "On finite groups with cyclic Sylow subgroups for all odd primes," *American Journal of Mathematics*, vol. 77 (1955), pp. 657-691.

[8] H. F. Tuan, "On groups whose order contains a prime number to the first power," *Annals of Mathematics*, vol. 45 (1944), pp. 110-140.

[9] H. Zassenhaus, *The Theory of Groups*, Chelsea, 1958.

THE FULL IMBEDDING THEOREM.

By Barry Mitchell.[1]

Introduction. The group valued imbedding theorem asserts that every small abelian category admits an exact, covariant imbedding into the category of abelian groups (Freyd [3], or Lubkin [6]). From this it follows that a statement concerning exactness and commutativity of a diagram in a general abelian category is true providing that it is true in the category of abelian groups. However, the non-fullness of the imbedding prevents one from handling statements which concern the existence of morphisms in a diagram. In this paper it will be proved that every small abelian category admits a full, exact, covariant imbedding into a category of $\mathcal{R}$-modules for some ring $\mathcal{R}$. Thus any statement involving the existence of morphisms in a finite diagram which is true in all categories of modules is true in the general abelian category. Composing such a full imbedding with the functor from modules to groups which forgets the $\mathcal{R}$-module structure gives a further proof of the group valued imbedding theorem.

In § 1 we prove a number of results dealing with abelian categories satisfying the axiom A. B 5 of Grothendieck [5]. Some of these results were stated in [5], but since there does not seem to be proofs of them in print, I have chosen to include such proofs here.

In § 2 the Eckmann-Schopf techniques for injective envelopes are handled in the setting of the A. B 5 category. Also it is shown here that every object in an A. B 5 category with a generator has an injective envelope. From this we reobtain the Grothendieck result that such a category has an injective cogenerator.

Module categories are characterized in § 3 as abelian categories (with infinite coproducts) which possess a small, projective generator. The theorem proved is actually a refinement of this statement, dealing with the case where the generator is not necessarily small.

A quick review of functor categories precedes the proof of the full imbedding theorem in § 4. The crucial theorem needed here states that the

---

Received April 27, 1964.

[1] This work has been partially supported by the National Science Foundation under Grant No. NSF GP1904.

category of kernel preserving functors from a small abelian category to the category of abelian groups is itself an abelian category (Freyd [3], or Gabriel [4]).

Finally in § 5 we point out a generalization of the important Lemma 4.1 and use it to construct a counter-example to the statement that injective objects in the functor category $(\mathcal{A}, \mathcal{B})$ $(\mathcal{A}, \mathcal{B}$ abelian categories) are cokernel preserving as functors (a true statement when $\mathcal{B}$ is the category of abelian groups).

A part of the material in this paper was written as a thesis under the helpful supervision of David Buchsbaum, who in particular suggested that there exist a proof of 2.9 along the lines given here. I have also profited from talks with Peter Freyd, some of whose ideas I have used in § 4.

1. Preliminaries.

Unless otherwise stated $\mathcal{A}$ will denote an abelian category with arbitrary coproducts (sometimes called direct sums. See [5], axiom A.B.3). The coproduct of a family of objects $\{A_i\}_{i \in I}$ will be denoted by $\bigoplus_{i \in I} A_i$, and $u_k\colon A_k \to \bigoplus_{i \in I} A_i$ will denote the k-th injection. If $A_i = A$ for all $i \in I$, then the coproduct will be denoted by ${}^I A$. If $\{A_i\}_{i \in I}$ is a family of subobjects of an object A (more correctly, if $\{A_i \to A\}$ is a family of monomorphisms), then the *union* $\bigcup_{i \in I} A_i$ is the image of the obvious morphism $\bigoplus_{i \in I} A_i \to A$. The union is the smallest subobject of A which contains each of the given subobjects. The *intersection* $A_1 \cap A_2$ of two subobjects $A_1, A_2 \subset A$ is the kernel of the composition $A_1 \to A \to A/A_2$. The intersection is the largest subobject of A which is contained in both A_1 and A_2 More generally, if $f\colon A \to B$ is any morphism and $B' \subset B$, then the *inverse image* $f^{-1}(B') \subset A$ is defined as the kernel of the composition $A \xrightarrow{f} B \to B/B'$. If $A' \subset B$ is the image of f then we have an exact sequence

$$0 \to f^{-1}(B') \to A \to A'/A' \cap B' \to 0.$$

The inverse image is the largest subobject of A which is carried into B' by f.

Consider a family of objects $\{A_i\}_{i \in I}$ where I is a directed set, and suppose that whenever $i \leqq j$ we have a morphism $\pi_{ij}\colon A_i \to A_j$ such that $\pi_{jk}\pi_{ij} = \pi_{ik}$ for $i \leqq j \leqq k$ and $\pi_{ii} = 1_{A_i}$ (the identity morphism on A_i) for all $i \in I$. Then we call $\{A_i, \pi\}$ a *direct system* over I. A *cocompatible* family for the direct system is a family of morphisms $\{A_i \to A\}_{i \in I}$ such that whenever $i \leqq j$ we have a commutative diagram

$$\begin{tikzcd}
A_i \arrow[d, "\pi_{ij}"'] \arrow[dr] & \\
& A \\
A_j \arrow[ur] &
\end{tikzcd}$$

The above family is called the *direct limit* for $\{A_i,\pi\}$ if given any other co-compatible family $\{A_i\to A'\}$ there is a unique morphism $A\to A'$ such that

$$A_i\to A\to A' = A_i\to A'$$

for all $i\in I$. In this case the morphism $A_i\to A$ is denoted by π_i, and the object A is denoted by $\underset{\substack{\longrightarrow\\ i\in I}}{\lim} > A_i$. In abelian categories with coproducts the direct limit can be taken as the quotient object

$$\bigoplus_{i\in I} A_i/\bigcup_{j<k}\operatorname{Image}(u_j - u_k\pi_{jk}). \tag{1}$$

The morphism π_i in this case is just the composition of u_i with the epimorphism from $\bigoplus A_i$ onto the quotient object (1). On the other hand, if $\{A,\pi_i\}$ and $\{A',\pi_i'\}$ are two direct limits for the same system, then one shows easily that the induced morphism $A\to A'$ is an isomorphism. Hence up to isomorphism direct limits are unique.

Let $\{A_i,\pi\}$ and $\{A_i',\pi'\}$ be two direct systems over I. A morphism $\{A_i,\pi\}\to\{A_i',\pi'\}$ of direct systems is a family of morphisms $f_i\colon A_i\to A_i'$ such that $\pi_{ij}'f_i = f_j\pi_{ij}$ whenever $i<j$. Such a morphism induces a morphism $f\colon A\to A'$ of the direct limits. The following proposition follows easily from the axiomatic descriptions of cokernels and direct limits.

PROPOSITION 1.1. *If* $\{A_i',\pi'\}\to\{A_i,\pi\}$ *and* $\{A_i,\pi\}\to\{A_i'',\pi''\}$ *are morphisms of direct systems such that* $A_i'\to A_i\to A_i''\to 0$ *is exact for all* i, *then the direct limit sequence* $A'\to A\to A''\to 0$ *is exact.* (*In other words direct limits preserve cokernels.*)

Let $\{A_i\}_{i\in I}$ be a family of subobjects of an object A. Define $i\leqq j$ if A_i precedes A_j as a subject of A. If I becomes directed under this ordering we say that the above family is a *direct family of subobjects.* Then taking π_{ij} as the inclusion morphism from A_i to A_j we obtain a direct system whose direct limit is not in general a subobject of A.

PROPOSITION 1.2. *The following properties of a category are equivalent.*

a) *The direct limit of a direct family of subobjects* $\{A_i\}$ *is* $\bigcup_{i\in I} A_i$.

b) *If* $\{A_i\}$ *is a direct family of subobjects of* A *and* B *is another subobject, then*

$$(\bigcup A_i) \cap B = \bigcup (A_i \cap B). \tag{2}$$

Proof. Assuming a), the direct limit of the family $\{A_i \cup B\}$ is

$$\bigcup (A_i \cup B) = (\bigcup A_i) \cup B.$$

Using 1.1 we then see that the direct limit of the system $\{A_i \cup B/B\}$ is $(\bigcup A_i) \cup B/B$. Hence again using 1.1 the family of exact sequences

$$0 \to A_i \cap B \to A_i \to A_i \cup B/B \to 0$$

gives us an exact direct limit sequence

$$0 \to \bigcup (A_i \cap B) \to \bigcup A_i \to (\bigcup A_i) \cup B/B \to 0 \tag{3}$$

But (3) is just another way of writing (2).

Now assume b). We have the morphism $u: D \to A$ induced from the direct limit D of the direct family of subobjects by the inclusion morphisms. Furthermore the image of u is $\bigcup A_i$. We wish to show that u is a monomorphism. Let K be the kernel of u, and denote the image of $A_i \to D$ by A_i'. Then $D = \bigcup A_i'$, and so by b) we have $K = \bigcup (A_i' \cap K)$. If K is not zero, then $A_i' \cap K$ is not zero for some i and consequently the inverse image of $K \cap A_i'$ in A_i is not zero. But this means that $A_i \to A$ has a non-zero kernel, contradicting the fact that it is a monomorphism. Hence $K = 0$, and so u is a monomorphism.

Proposition 1.2 proves the equivalence of two formulations of the Grothendieck axiom A.B.5 [5].

The proof of the following lemma is left to the reader.

LEMMA 1.3. *Let $\{A_i\}$ be a direct family of subobjects of A. Then the direct limit of the corresponding family of quotient objects is $A/\bigcup A_i$.*

PROPOSITION 1.4. A.B.5 *is equivalent to the condition*

$$f^{-1}(\bigcup A_i) = \bigcup f^{-1}(A_i) \tag{4}$$

relative to any morphism $f: B \to A$ and any direct family of subobjects $\{A_i\}$ of A.

Proof. Taking f to be the inclusion of a subobject $B \subset A$, we obtain equation (2).

Conversely, suppose that A.B.5 is satisfied, and let f be any morphism. Let B' be the image of f. Then we have an exact sequence

$$0 \to f^{-1}(A_i) \to B \to B'/A_i \cap B' \to 0.$$

Passing to direct limits and using 1.3 we obtain an exact sequence

$$(5) \qquad 0 \to \bigcup f^{-1}(A_i) \to B \to B'/\bigcup\, (A_i \cap B') \to 0.$$

But by A.B.5 we have $\bigcup\, (A_i \cap B') = (\bigcup A_i) \cap B'$. Hence (5) is just another way of expressing (4).

PROPOSITION 1.5. *Consider a direct system* $\{A_i, \pi\}_{i\in I}$ *in an* A.B.5 *category. Denote* $K_{ij} = \text{kernel}\, \pi_{ij}$ *for* $i \leqq j$, *and denote* $K_i = \text{kernel}\, \pi_i$. *Then for any index* k *we have*

$$K_k = \bigcup_{k \leqq p} K_{kp}.$$

Proof. Let R be the set of ordered pairs of indices (i, j) such that $i < j$. For any subset S of R let

$$A_S = \bigcup_{(i,j)\in S} \text{Image}(u_i - u_j\pi_{ij}) \subset \bigoplus_{i\in I} A_i.$$

Then by (1) we have $\underset{i\in I}{\underrightarrow{\lim}} > A_i = \bigoplus_{i\in I} A_i/A_R$. Also we have $A_R = \bigcup_F A_F$ where F runs through all finite subsets of R. Therefore by 1.4 we can write

$$K_k = u_k^{-1}(A_R) = \bigcup_F u_k^{-1}(A_F).$$

Hence the proposition will be proved if we can show that for each F there is an index p such that $u_k^{-1}(A_F) \subset K_{kp}$. We take for p any index which follows k and all indices which appear in either first or second position of a member of F. Now $u_k^{-1}(A_F)$ is the kernel of the composition

$$(6) \qquad A_k \to \bigoplus_{i\in I} A_i \to \bigoplus_{i\in I} A_i/A_F.$$

Define $f\colon \bigoplus_{i\in I} A_i \to A_p$ by taking $fu_i = \pi_{ip}$ for $i \leqq p$ and $fu_i = 0$ otherwise. Then f extends π_{kp} in (6). Moreover for $(i, j) \in F$ we have $f(u_i - u_j\pi_{ij}) = 0$. Hence f factors through $\bigoplus_{i\in I} A_i/A_F$ in (6). From this it follows that $u_k^{-1}(A_F) \subset K_{kp}$.

PROPOSITION 1.6. A.B.5 *is equivalent with the assertion that direct limits are exact.*

Proof. If direct limits are exact, then the limit morphism of direct family of monomorphisms $A_i \to A$, which can be regarded as a morphism from $\{A_i\}$ to the constant direct system which as A at every index, is a monomorphism. Hence by 1.2 A.B.5 is satisfied.

Conversely suppose that A.B.5 is satisfied. Since in any case direct limits preserve cokernels (1.1) it suffices to show that the limit morphism $f\colon A \to B$ of a family of monomorphisms $f_i\colon A_i \to B_i$ defining a morphism from the direct system $\{A_i, \pi\}_{i \in I}$ to the direct system $\{B_i; \eta\}_{i \in I}$ is a monomorphism. Consider the commutative diagram

$$\begin{tikzcd}
 & A_i \arrow[r, "f_i"] \arrow[d, "\pi_i"] & B_i \arrow[d, "\eta_i"] \\
K \arrow[r] & A \arrow[r, "f"] & B
\end{tikzcd}$$

where K is the kernel of f. Denoting $A_i' = \text{image}\,\pi_i$ we have $A = \bigcup A_i'$. By A.B.5 we then have $K = \bigcup (A_i' \cap K)$. Suppose $K \neq 0$. Then $A_i' \cap K \neq 0$ for some i. Letting $M = \pi_i^{-1}(A_i' \cap K)$ we have $\pi_i \mid M \neq 0$ whereas $\eta_i f_i \mid M = 0$. Then $M \subset \bigcup_{i<j} L_{ij}$ by 1.5, where $L_{ij} = \text{kernel}\,\eta_{ij}$, and so again using A.B.5 we can write $M = \bigcup_{i<j} (L_{ij} \cap \mathrm{M}$. But $L_{ij} \cap M$, considered as a subobject of A_i, must go into zero under π_{ij} since $A_j \to B_j$ is a monomorphism. Hence $L_{ij} \cap M$ goes into zero under π_i for all j, and consequently $\pi_i \mid M = 0$. This contradiction proves that $K = 0$, and so f is a monomorphism.

2. Injective envelopes. Following Eckmann and Schopf [2] we define an *essential extension* of A' to be a monomorphism $A' \to A$ such that every non-zero subobject of A intersects A' non-trivially. In the category $\mathcal{G}^{\mathcal{R}}$ of right $\mathcal{R}$-modules $A' \subset A$ is an essential extension if and only if for each $a \neq 0$ in A there is an $r \in \mathcal{R}$ such that $0 \neq ar \in A'$.

LEMMA 2.1. *Let $u\colon A' \to A$ be a monomorphism. Then the following are equivalent.*

a) *u is an essential extension.*

b) *Every morphism $f\colon A \to B$ such that fu is a monomorphism is a monomorphism.*

Proof. a) $\Longrightarrow$ b) If *f is not a monomorphism, then* $K = \text{kernel}\, f \neq 0$, hence $K \cap A' \neq 0$. But $K \cap A' \subset \text{kernel}\,(fu)$. Therefore fu is not a monomorphism.

b) $\Longrightarrow$ a) If u is not an essential extension, then there is a non-zero subobject B of A such that $B \cap A' = 0$. But this means that the composition $A' \to A \to A/B$ is a monomorphism. Therefore by b), $A \to A/B$ is a monomorphism, which is not so.

Given morphisms $u: A' \to A$ and $f: A' \to B$, we form the quotient object

$$P = A \oplus B/\text{Image}(u_A u - u_B f)$$

where u_A and u_B denote the injections into the coproduct. Letting g and v denote the compositions of u_A and u_B respectively with the natural epimorphism $A \oplus B \to P$, it is clear that we have a commutative diagram

$$\begin{tikzcd} A' \arrow[r, "u"] \arrow[d, "f"'] & A \arrow[d, "g"] \\ B \arrow[r, "v"] & P \end{tikzcd} \tag{1}$$

The diagram (1) is called the *pushout* of the two morphisms u and f.

LEMMA 2.2. *If in the pushout diagram* (1) *the morphism u is a monomorphism, then so is v.*

Proof. Let $K = \text{kernel}\, v$, and suppose $K \neq 0$. Denote $h = u_A u - u_B f$. Then $u_B(K) \subset \text{Image}\, h$. Therefore $K' = h^{-1}(u_B(K)) \neq 0$. Letting p_A be the projection $A \oplus B \to A$, we have $p_A h = p_A(u_A u - u_B f) = u$. Since u is a monomorphism this means that $p_A h(K') \neq 0$. But

$$p_A h(K') = p_A h(h^{-1}(u_B(K))) \subset p_A u_B(K) = 0.$$

This contradiction proves the lemma.

For the next lemma we need the following definitions. We call a category *locally small* if every object has only a set of (equivalence classes of) subobjects. We say that a morphism $u_1: A_1 \to A$ is a *retract* if there is a morphism $p_1: A \to A_1$ such that $p_1 u_1 = 1_{A_1}$. In an abelian category the morphism u_1 is a retract if and only if there is a subobject $A_2 \subset A$ such that A is the coproduct of A_1 and A_2 with u_1 as the first injection. Also for abelian categories a union $A_1 \cup A_2$ of two subobjects is a coproduct if and only if $A_1 \cap A_2 = 0$.

LEMMA 2.3. *Let Q be an object in a locally small* A.B.5 *category. Then the following are equivalent.*

a) *Q is injective.*

b) *Q admits no proper essential extensions.*

c) *Q is a retract of every object which contains it.*

Proof. a) $\Longrightarrow$ b) Suppose that $u\colon Q \to A$ is an essential extension. Since Q is injective we can find a morphism $p\colon A \to Q$ such that $pu = 1_Q$. That is, Q is a retract of A. Letting Q' be such that $A = Q \oplus Q'$ we have $Q \cap Q' = 0$. Since u is an essential extension this means that $Q' = 0$. Hence u is an isomorphism.

b) $\Longrightarrow$ c) Suppose that Q admits no proper essential extensions. Let Q be a subobject of A but not a retract of A. Let $\mathcal{B}$ be the set of subobjects of A which intersect Q trivially. If $\{A_i\}$ is a linearly ordered subset of $\mathcal{B}$, then by A. B. 5 we have

$$0 = \bigcup (A_i \cap Q) = (\bigcup A_i) \cap Q.$$

This shows that $\mathcal{B}$ is inductive and so has a maximal element A'. Now $Q = Q \oplus A'/A' \subset A/A'$. Since $Q \oplus A' \neq A$, A/A' is a proper extension of Q, hence is not essential. Let $\bar{B}$ be a non-zero subobject of A/A' which intersects Q trivially. Then $\bar{B}$ is of the form B/A', where B is a subobject of A and A' is a proper subobject of B. Furthermore since $\bar{B} \cap Q = 0$ and $A' \cap Q = 0$, we must have $B \cap Q = 0$. But this contradicts the maximality of A'.

c) $\Longrightarrow$ a). Consider a monomorphism $u\colon A' \to A$ and a morphism $f\colon A' \to Q$. We wish to find $h\colon A \to Q$ such that $hu = f$. We form the pushout diagram

$$\begin{tikzcd}
A' \arrow[r, "u"] \arrow[d, "f"'] & A \arrow[d, "g"] \\
Q \arrow[r, "v"] & P
\end{tikzcd}$$

where by 2.2 we know that v is a monomorphism. Then c) says that we can find $w\colon P \to Q$ such that $wv = 1_Q$. Hence we can take $h = wg$.

The following lemma is left to the reader.

Lemma 2.4. *If $u\colon A \to B$ and $v\colon B \to C$ are monomorphisms, then vu is an essential extension if and only if both u and v are essential extensions.*

Lemma 2.5. *Suppose that $\{A_i\}$ is a direct family of subobjects of A in an* A. B. 5 *category and suppose that each is an essential extension of a subobject A' of A. Then $\bigcup A_i$ is an essential extension of A'.*

Proof. Let B be a non-zero subobject of $\bigcup A_i$. By A. B. 5 we have $B = \bigcup (A_i \cap B)$ and so $A_i \cap B \neq 0$ for some i. Hence since A_i is an essential extension of A', we have $A_i \cap B \cap A' \neq 0$, and so A' intersects B non-trivially.

A functor $T\colon \mathcal{A} \to \mathcal{B}$ is called *faithful* if for every pair of objects A_1, A_2 in $\mathcal{A}$ the induced function

$$(2) \qquad [A_1, A_2]_{\mathcal{A}} \to [T(A_1), T(A_2)]_{\mathcal{B}}$$

is univalent, where $[A_1, A_2]$ denotes the group of morphisms from A_1 to A_2. A faithful functor which takes distinct objects into distinct objects is called an *imbedding*. If on the other hand (2) is always onto, then T is called *full*.

Given an object $U \in \mathcal{A}$, consider the abelian group valued functor $T\colon \mathcal{A} \to \mathcal{G}$ defined by $T(A) = [U, A]$. Then U is called a *generator* for $\mathcal{A}$ if T is an imbedding. That is, U is a generator if for any non-zero morphism $f\colon A \to B$ there is a morphism $g\colon U \to A$ such that $fg \neq 0$. Equivalently U is a generator if whenever A' is a proper subobject of A there is a morphism $U \to A$ whose image is not in A'. From this we see that a subobject of A is determined (up to isomorphism of subobjects) by the morphisms $U \to A$ which factor through it. Since $[U, A]$ is a set for all A, it follows that a category with a generator is locally small.

The following lemma is immediate from the definition of generator.

LEMMA 2.6. *U is a generator for $\mathcal{A}$ if and only if for every object $A \in \mathcal{A}$ the morphism ${}^{[U,A]}U \to A$ whose g-th coordinate is g is an epimorphism.*

If U is a generator for $\mathcal{A}$, then an object will be called *finitely generated* (with respect to U) if it is a quotient of an object of the form ${}^{I}U$ for some finite set I. Clearly the union of two finitely generated subobjects is also finitely generated, and consequently the finitely generated subobjects of an object A comprise a direct family of subobjects. If $\mathcal{A}$ is an A.B.5 category it follows from 1.2 that the direct limit of the family is A.

If U is projective, then U is a generator if and only if $[U, A] \neq 0$ for every non-zero object A in $\mathcal{A}$. From 2.6 it follows that a category with a projective generator has projective resolutions. In the category of right (resp. left) $\mathcal{R}$-modules $\mathcal{G}^{\mathcal{R}}$ (resp. ${}^{\mathcal{R}}\mathcal{G}$) a projective generator is $\mathcal{R}$ considered as a right (resp. left) module over itself.

Dually C is called a *cogenerator* if for any non-zero morphism $f\colon B \to A$ there is a morphism $g\colon A \to C$ such that $gf \neq 0$. If C is injective then it is a cogenerator if and only if $[A, C] \neq 0$ for every non-zero object A. It is well known that an abelian group is injective if and only if it is divisible (see for example [1], page 8, Theorem 3.2). Consequently if $\boldsymbol{Q}$ denotes the additive group of rational numbers and $\boldsymbol{Z}$ denotes the integers, then $\boldsymbol{Q}/\boldsymbol{Z}$ is injective. Using this it is not difficult to construct a non-zero morphism of groups $G \to \boldsymbol{Q/Z}$ for every non-zero abelian group G. Hence $\boldsymbol{Q/Z}$ is an injective cogenerator for the category of abelian groups $\mathfrak{G}$. Then using the obvious natural equivalence of functors of the right $\mathfrak{R}$-module A

$$[A, [\mathfrak{R}, \boldsymbol{Q/Z}]_G]_{\mathfrak{R}} \approx [A, \boldsymbol{Q/Z}]_G$$

we see that $[\mathfrak{R}, \boldsymbol{Q/Z}]_G$ considered as a right $\mathfrak{R}$-module is an injective cogenerator for $\mathfrak{G}^{\mathfrak{R}}$. Likewise $[\mathfrak{R}, Q/Z]_G$ considered as a left $\mathfrak{R}$-module is an injective cogenerator for ${}^{\mathfrak{R}}\mathfrak{G}$.

Given $U \in \mathcal{A}$, let $\mathfrak{R}$ denote the ring of endomorphisms of U. Then for $A \in \mathcal{A}$ the group $[U, A]$ can be considered as a right $\mathfrak{R}$-module by defining the operation of the ring element $r \in [U, U]$ on the group element $a \in [U, A]$ as the composition $ar \in [U, A]$. Hence the functor T defined by $T(A) = [U, A]$ may be considered as having values in the category of right $\mathfrak{R}$-modules.

LEMMA 2.7. *If U is a generator, then $T \colon \mathcal{A} \to \mathfrak{G}^{\mathfrak{R}}$ preserves essential extensions.*

Proof. Suppose that $u \colon A' \to A$ is an essential extension in $\mathcal{A}$, and let f be a non-zero element of $T(A) = [U, A]$. We must find $r \in \mathfrak{R} = [U, U]$ such that $0 \neq fr$ and $fr \in [U, A']$ (or more precisely, $\operatorname{Image}(fr) \subset A'$). Since u is an essential extension we have $A' \cap \operatorname{Image} f \neq 0$. Therefore we have a commutative diagram

$$\begin{tikzcd}
f^{-1}(A') \arrow[r, "f'"] \arrow[d, "v"'] & A' \arrow[d, "u"] \\
U \arrow[r, "f"] & B
\end{tikzcd}$$

where $f' \neq 0$. Let $g \colon U \to f^{-1}(A')$ be such that $f'g \neq 0$. Then we can take vg as the required r.

If $u \colon A \to Q$ is an essential extension and Q is injective, then u is called an *injective envelope* for A.

PROPOSITION 2.8. *If $u \colon A \to Q$ and $u_1 \colon A \to Q_1$ are injective envelopes for A, then there is an isomorphism $Q \approx Q_1$ (not necessarily unique) such that the diagram*

$$\begin{tikzcd}
 & A \arrow[dl, "u"'] \arrow[dr, "u_1"] & \\
Q \arrow[rr, phantom, "\approx"] & & Q_1
\end{tikzcd}$$

is commutative.

Proof. Let $\theta\colon Q \to Q_1$ be such that $u_1 = \theta u$ (injectivity of Q_1). Then θ is a monomorphism (essentiality of u) and so by 2.4 θ is an essential extension (essentiality of u_1). Therefore θ is an isomorphism (injectivity of Q).

THEOREM 2.9. *An* A.B.5 *category with a generator has injective envelopes for each of its objects.*

Proof. Let T and $\mathfrak{R}$ be as above, and let $T(A) \to M$ be a monomorphism of right $\mathfrak{R}$-modules with M injective. Consider the class $\mathcal{B}$ of all triples (B, u, f) where $u\colon A \to B$ is an essential extension and $f\colon T(B) \to M$ is such that the diagram

$$\begin{tikzcd} T(A) \arrow[r, "T(u)"] \arrow[dr] & T(B) \arrow[d, "f"] \\ & M \end{tikzcd}$$

is commutative. By 2.7 $T(u)$ is an essential extension and so by 2.1 f must be a monomorphism. We order $\mathcal{B}$ by defining $(B, u, f) \leqq (B', u', f')$ if there is a morphism $v\colon B \to B'$ such that $vu = u'$ and such that the diagram

$$\begin{tikzcd} T(B) \arrow[r, "T(v)"] \arrow[dr, "f"'] & T(B') \arrow[d, "f'"] \\ & M \end{tikzcd}$$

is commutative. Such a v must be a monomorphism since u is an essential extension. Furthermore it must be unique. For if v' were another then $fT(v) = fT(v')$, and so since f is a monomorphism and T is faithful we have $v = v'$. It follows that if also $(B', u', f') \leqq (B, u, f)$ then v is an isomorphism. In this case we shall say that (B, u, f) and (B', u', f') are equivalent. Now since $f\colon T(B) \to M$ is a monomorphism we see that the cardinality of $T(B) = [U, B]$ is less than or equal to the cardinality of M. Consequently by 2.6 B can be written as a quotient of ${}^{M}U$. But $\mathfrak{A}$ having a generator is locally small, and consequently there is no more than a set of quotient objects of ${}^{M}U$. It follows that the class $\mathcal{B}_0$ of equivalence classes in $\mathcal{B}$ is a set. Let $\{(B_i, u_i, f_i)\}$ represent a linearly ordered subset of $\mathcal{B}_0$. Then using uniqueness of the morphism v, we se that A together with the B_i's form a direct system in which all the morphisms are monomorphisms. Therefore using 1.5 we can regard this direct system as a directed family of subobjects of the direct limit B_0. Moreover by 2.5 the inclusion $u_0: A \to B_0$ is an essential extension. Now $T(A)$ together with the $T(B_i)$'s form a direct family of submodules of $T(B_0)$, and so the compatible family of morphisms $f_i: T(B_i) \to M$ defines a limit morphism $f: \cup T(B_i) \to M$ which can be extended to a morphism $f_0: T(B_0) \to M$ by injectivity of M. Thus the class of (B_0, u_0, f_0) dominates every (B_i, u_i, f_i) and so we have shown that $\mathcal{B}_0$ is inductive. Applying Zorn's lemma we have a maximal member (Q, w, h). If Q were not injective, then by 2.3 we could find a proper essential extension $v_1: Q \to Q_1$ and by 2.4 $v_1 w$ is an essential extension. Also by injectivity of M the morphism h could be extended to a morphism $h_1: T(Q_1) \to M$. But this contradicts the maximality of (Q, w, h). Therefore Q is injective, and so w is an injective envelope for A.

PROPOSITION 2.10. *If $\mathcal{A}$ has injective resolutions and a generator, then $\mathcal{A}$ has an injective cogenerator.*

Proof. Consider a monomorphism $\bigoplus_{V \subset U} U/V \to Q$ where U is the generator, V runs through all subobjects of U, and Q is injective. We show that Q is a cogenerator. It suffices to show that for any $A \neq 0$ there is a non-zero morphism $A \to Q$. Since U is a generator there is a non-zero morphism $U \to A$. The image of this morphism is U/V_0 for some subobject $V_0 \neq U$. Then by injectivity of Q the non-zero composition $U/V_0 \to \oplus U/V \to Q$ can be extended to a morphism $A \to Q$ as required.

Combining 2.9 and 2.10 we obtain

COROLLARY 2.11. *An* A.B.5 *category with a generator has an injective cogenerator.*

The following proposition shows that injective envelopes may exist even when $\mathcal{A}$ does not have a generator.

PROPOSITION 2.12. *Let A be an object in a locally small* A.B.5 *category and suppose that A is a subobject of an injective Q. Then A has an injective envelope.*

Proof. Consider the set $\mathcal{B}$ of all subobjects of Q which are essential extensions of A. Then by 2.5 $\mathcal{B}$ is inductive, and so let Q_0 be maximal in $\mathcal{B}$. If Q_0 is not injective then by 2.3 there is a proper essential extension $u: Q_0 \to B$. Since Q is injective there is a morphism $v: B \to Q$ such that vu is the inclusion $Q_0 \subset Q$. But since u is essential v is a monomorphism. Hence using 2.4 we see that B is an essential extension of A in Q, contradicting the maximality of Q_0. Therefore Q_0 is injective.

3. Characterization of module categories. A functor $T\colon \mathcal{A} \to \mathcal{B}$ ($\mathcal{A}$ and $\mathcal{B}$ not necesssarily abelian) is called an *equivalence* if it is a full imbedding such that every object $B \in \mathcal{B}$ is isomorphic to $T(A)$ for some $A \in \mathcal{A}$. It is not difficult to prove that T is an equivalence if and only if there is a functor $S\colon \mathcal{B} \to \mathcal{A}$ such that ST and TS are naturally equivalent to the identity functors on $\mathcal{A}$ and $\mathcal{B}$ respectively. Furthermore S and T preserve any situation in a category which can be described in terms of universal mapping properties (kernels, products, intersections, etc.). In other words if $T\colon \mathcal{A} \to \mathcal{B}$ is an equivalence (or if $\mathcal{A}$ and $\mathcal{B}$ are equivalent as we shall say) then for all practical purposes $\mathcal{A}$ and $\mathcal{B}$ are the same.

An object U in an abelian category is called *small* if every morphism from U into a coproduct $U \to \bigoplus_{i \in I} A_i$ factors as $U \to \bigoplus_{i \in J} A_i \to \bigoplus_{i \in I} A_i$ where J is a finite subset of I and the morphism between the coproducts is the one which preserves injections. Denoting $T(A) = [U, A]$ one can easily show that U is small if and only if T preserves coproducts. In the category $\mathcal{G}^{\mathcal{R}}$ the ring $\mathcal{R}$ considered as a right module is a small projective generator. On the other hand if $\mathcal{A}$ is a category with a small projective generator U, then the following theorem shows that $\mathcal{A}$ is equivalent to a category of modules. We also give an analysis of the situation where U is not necessarily small. This will be needed in the sequel.

THEOREM 3.1. *Let $\mathcal{A}$ be a category with a projective generator U and let $\mathcal{R}$ denote the ring of endomorphisms of U. Then the functor $T\colon \mathcal{A} \to \mathcal{G}^{\mathcal{R}}$ defined by $T(A) = [U, A]$ is an exact imbedding. The induced function*

$$[A, B] \to [T(A), T(B)]$$

is onto whenever A is finitely generated. If either U is small or $\mathcal{A}$ is A.B.5, *then T is full. T is an equivalence if and only if U is small.*

Proof. The conditions that U be projective and a generator are equivalent to T being exact and an imbedding respectively. We must show that if A is finitely generated, then every morphism of $\mathcal{R}$-modules $\Phi\colon [U, A] \to [U, B]$ can be realized as $T(\phi)$ for some $\phi \in [A, B]$. That is, we must find ϕ such that $\Phi(f) = \phi f$ for all $f \in [U, A]$. Suppose first that $U = A$. Then we have

$$\Phi(f) = \Phi(1_U) = \Phi(1_U) f,$$

the last equality being true since Φ is $\mathcal{R}$-linear. Hence in this case we may take $\phi = \Phi(1_U)$.

Now assume that we have an exact sequence

$$0 \to K \xrightarrow{\lambda} {}^{I}U \xrightarrow{\pi} A \to 0$$

where I is finite. Consider the composition

$$[U,U] \xrightarrow{T(u_k)} [U,{}^{I}U] \xrightarrow{T(\pi)} [U,A] \xrightarrow{\Phi} [U,B]$$

and call it Φ_k. Then we have shown above that we can find $\phi_k\colon U \to B$ such that $\Phi_k(f) = \phi_k f$ for all $f \in [U,U]$. The ϕ_k's then define a morphism $\bar{\phi}\colon {}^{I}U \to B$. If we denote $\bar{\Phi}$ as the composition $\Phi T(\pi)$ and denote the k-th projection from the coproduct by p_k, then we have

$$\begin{aligned} \bar{\Phi}(f) &= \Phi(\pi f) = \Phi(\pi \sum_{i\in I} u_i p_i f) \\ &= \sum_{i\in I} \Phi(\pi u_i p_i f) = \sum_{i\in I} \Phi_i(p_i f) \\ &= \sum_{i\in I} \phi_i p_i f = \bar{\phi} f. \end{aligned} \tag{1}$$

We show that $\bar{\phi}\lambda = 0$. If $\bar{\phi}\lambda \neq 0$ then there is a morphism $\alpha\colon U \to K$ such that $\bar{\phi}\lambda\alpha \neq 0$. But

$$\bar{\phi}\lambda\alpha = \bar{\Phi}(\lambda\alpha) = \Phi(\pi\lambda\alpha) = 0,$$

a contradiction. Therefore $\bar{\phi}\lambda = 0$, and so $\bar{\phi}$ induces $\phi\colon A \to B$ such that $\phi\pi = \bar{\phi}$. Now take any $f\colon U \to A$. Then since U is projective we may write $f = \pi g$ where $g\colon U \to {}^{I}U$. Hence using (1) we have

$$\Phi(f) = \Phi(\pi g) = \bar{\Phi}(g) = \bar{\phi} g = \phi\pi g = \phi f.$$

If U is small, then the above goes through even when A is not finitely generated. For if $f\colon U \to {}^{I}U$ (where I is not necessarily finite) then we may assume that f factors through ${}^{J}U$ where J is a finite subset of I, and then in (1) we can write $f = \sum_{i\in J} u_i p_i f$.

Now suppose that $\mathcal{A}$ is A.B.5 (but that U is not necessarily small). Given an object A we can write $A = \varinjlim_{i\in I} A_i$ where $\{A_i\}_{i\in I}$ is the set of all finitely generated subobjects. For each $i \in I$ define Φ_i as the composition

$$[U,A_i] \to [U,A] \xrightarrow{\Phi} [U,B].$$

Then each Φ_i is determined by a $\phi_i\colon A_i \to B$. Furthermore the ϕ_i's are unique since T is an imbedding. From this it follows that $\{\phi_i\}$ is a cocompatible system and so there is an induced morphism $\phi\colon A \to B$. Now if $f \in [U, A]$ then image f is an A_i for some i, and so we can write $f = \pi_i g$ where $g\colon U \to A_i$ and π_i is the inclusion $A_i \to A$. Then

$$\Phi(f) = \Phi(\pi_i g) = \Phi_i(g) = \phi_i g = \phi \pi_i g = \phi f$$

and so ϕ is the required morphism.

Finally, if T is an equivalence, then U is small since its image $\mathfrak{R}$ is small in $\mathcal{G}^{\mathfrak{R}}$. Conversely suppose that U is small and let M be any right $\mathfrak{R}$-module. We may write an exact sequence of $\mathfrak{R}$-modules:

$${}^I\mathfrak{R} \xrightarrow{\Phi} {}^J\mathfrak{R} \to M \to 0.$$

Since U is small, T commutes with coproducts, and so $T({}^IU) = {}^I\mathfrak{R}$ and $T({}^JU) = {}^J\mathfrak{R}$. Since T is full we have $\Phi = T(\phi)$ for some ϕ. Then since T is exact we have $M \approx T(C)$ where C is the cokernel of ϕ. This proves that T is an equivalence.

As an immediate application of 3.1, let $U = {}^n\mathfrak{R}$ in the category $\mathcal{G}^{\mathfrak{R}}$ where n is any positive integer. Then U is a small projective generator. The ring of endomorphisms of U is isomorphic to the ring $M_n(\mathfrak{R})$ of all $n \times n$ matrices over $\mathfrak{R}$. Consequently for any integer n the categories $\mathcal{G}^{\mathfrak{R}}$ and $\mathcal{G}^{M_n(\mathfrak{R})}$ are equivalent. A consideration of opposite rings shows also that ${}^{\mathfrak{R}}\mathcal{G}$ and ${}^{M_n(\mathfrak{R})}\mathcal{G}$ are equivalent. This will be considered as a part of a more general picture in a forthcoming paper.

As a further application we can show that the property of freeness for a module is not preserved under equivalences of categories. Let $\mathfrak{R}$ be a commutative ring, and take $U = \mathfrak{R} \oplus \mathfrak{R}$. Under the functor T of 3.1 the $\mathfrak{R}$-module $\mathfrak{R}$ corresponds to the Λ-module $[\mathfrak{R} \oplus \mathfrak{R}, \mathfrak{R}]$ where $\Lambda = [\mathfrak{R} \oplus \mathfrak{R}, \mathfrak{R} \oplus \mathfrak{R}]$. We show that every non-zero element of $[\mathfrak{R} \oplus \mathfrak{R}, \mathfrak{R}]$ has torsion, so that in particular $[\mathfrak{R} \oplus \mathfrak{R}, \mathfrak{R}]$ cannot be free. Let f be any such element. Suppose $f(1,0) = r$ and $f(0,1) = s$. Then r and s are not both zero. Hence $(s, -r)$ is a non-zero element of $\mathfrak{R} \oplus \mathfrak{R}$. Letting $g\colon \mathfrak{R} \oplus \mathfrak{R} \to \mathfrak{R} \oplus \mathfrak{R}$ be such that $g(1,0) = g(0,1) = (s, -r)$ we see that g is a non-zero element of Λ such that $fg = 0$.

4. The full imbedding theorem. Let $\mathcal{A}$ and $\mathcal{B}$ be any additive categories and supose that $\mathcal{A}$ is small (that is, the objects of $\mathcal{A}$ form a set). Then using natural transformations as morphisms the additive covariant functors from $\mathcal{A}$ to $\mathcal{B}$ form an additive category which we denote by $(\mathcal{A}, \mathcal{B})$. If $\mathcal{B}$ has kernels, then the kernel $K \to S$ of a morphism $S \to T$ in $(\mathcal{A}, \mathcal{B})$

is such that $K(A) \to S(A)$ is the kernel of $S(A) \to T(A)$ in $\mathcal{B}$ for all $A \in \mathcal{A}$. Likewise if $\mathcal{B}$ has coproducts then the coproduct $\bigoplus_{i \in I} T_i$ in $(\mathcal{A}, \mathcal{B})$ is given "pointwise" by the relation

$$(\bigoplus_{i \in I} T_i)(A) = \bigoplus_{i \in I} T_i(A).$$

By such pointwise considerations it is not hard to show that if $\mathcal{B}$ is abelian then so is $(\mathcal{A}, \mathcal{B})$ for any small additive category $\mathcal{A}$, and that if $\mathcal{B}$ is A.B.5, then so is $(\mathcal{A}, \mathcal{B})$. In particular this is true when $\mathcal{B} = \mathcal{G}$. In this case we have for each $A \in \mathcal{A}$ the object $H^A \in (\mathcal{A}, \mathcal{B})$ defined by $H^A(A') = [A, A']$. Clearly a morphism $A_1 \to A_2$ induces a natural transformation $H^{A_2} \to H^{A_1}$ and in this way we can regard $H\colon \to (\mathcal{A}, \mathcal{G})$ as a contravariant additive functor which takes cokernels into kernels. We have the following lemma which has become standard.

LEMMA 4.1. *Let $\mathcal{A}$ be any additive category and let $T\colon \mathcal{A} \to \mathcal{G}$ be a covariant additive functor. Then for any $A \in \mathcal{A}$ we have an isomorphism of groups*

(1) $$\theta\colon [H^A, T] \to T(A).$$

which is natural in both T and A.

Proof. We give only the definition of θ and its inverse and leave the details to the reader. We define θ by the rule $\theta(\eta) = \eta_A(1_A)$ for a natural transformation $\eta\colon H^A \to T$. The inverse θ' is given by $\theta'(x)_{A'}(f) = T(f)(x)$ for $x \in T(A)$ and $f \in [A, A']$.

It follows from 4.1 that for each $A \in \mathcal{A}$ ($\mathcal{A}$ small) the functor H^A is projective in $(\mathcal{A}, \mathcal{G})$. For if $S \to T$ is an epimorphism in $(\mathcal{A}, \mathcal{G})$ then the induced morphism $[H^A, S] \to [H^A, T]$ is equivalent to $S(A) \to T(A)$, and the latter is an epimorphism in $\mathcal{G}$ since the epimorphisms in $(\mathcal{A}, \mathcal{G})$ are characterized as being pointwise epimorphisms. Since a coproduct of projectives is again projective we see that $\bigoplus_{A \in \mathcal{A}} H^A$ is projective. Then for any $T\mathcal{A}\ (\mathcal{A}, \mathcal{G})$ we have

$$[\bigoplus_{A \in \mathcal{A}} H^A, T] = \mathbf{X}_{A \in \mathcal{A}} [H^A, T] = \mathbf{X}_{A \in \mathcal{A}} T(A).$$

Now $\mathbf{X}\, T(A)$ is zero only if T is zero. Consequently $\bigoplus_{A \in \mathcal{A}} H^A$ is a generator for $(\mathcal{A}, \mathcal{G})$.

Now let $\mathcal{A}$ be a small abelian category, and let $\mathcal{L}(\mathcal{A}, \mathcal{G})$ denote the full subcategory of $(\mathcal{A}, \mathcal{G})$ consisting of all kernel preserving functors. Since the kernel in $(\mathcal{A}, \mathcal{B})$ of a morphism of two kernel preserving functors is again kernel preserving, it follows that $\mathcal{L}(\mathcal{A}, \mathcal{B})$ has kernels. However, it is not obvious what the cokernel of a morphism in $\mathcal{L}(\mathcal{A}, \mathcal{B})$ should be, or even that such exists. Nevertheless, Freyd [3] has proved the following theorem, which we state without proof. The result also follows from work done previously by Gabriel [4].

THEOREM 4.2. *If* $\mathcal{A}$ *is any small category then* $\mathcal{L}(\mathcal{A}, \mathcal{B})$ *is abelian.*

It follows from the fact that coproducts in $\mathcal{B}$ preserve kernels that the coproduct in $(\mathcal{A}, \mathcal{B})$ of a family in $\mathcal{L}(\mathcal{A}, \mathcal{B})$ is also in $\mathcal{L}(\mathcal{A}, \mathcal{B})$. Consequently $\bigoplus_{A \in} H^A$ is in $\mathcal{L}(\mathcal{A}, \mathcal{B})$, and so it is a generator for $\mathcal{L}(\mathcal{A}, \mathcal{B})$ (but not necessarily projective). Also since $\mathcal{B}$ is A.B. 5 we see from 1.6 that the direct limit in $(\mathcal{A}, \mathcal{B})$ of a direct system in $\mathcal{L}(\mathcal{A}, \mathcal{B})$ is again in $\mathcal{L}(\mathcal{A}, \mathcal{B})$. Then again using the fact that $\mathcal{B}$ is A.B. 5, a pointwise consideration shows that direct limits in $\mathcal{L}(\mathcal{A}, \mathcal{B})$ are kernel preserving. Consequently $\mathcal{L}(\mathcal{A}, \mathcal{B})$ is A.B. 5. Therefore by 2.11 $\mathcal{L}(\mathcal{A}, \mathcal{B})$ has an injective cogenerator.

LEMMA 4.3. *If* $\mathcal{A}$ *is a small abelian category, then the functor* H: $\mathcal{A} \to \mathcal{L}(\mathcal{A}, \mathcal{B})$ *which assigns to the object* A *the functor* H^A *is an exact, contravariant, full imbedding.*

Proof. Replacing T by H^B in 4.1 we obtain $[H^A, H^B] \approx H^B(A) = [B, A]$. This shows that H is a contravariant full imbedding. Since H takes cokernels into kernels, to show exactness it suffices to show that if $B \to A$ is a monomorphism in $\mathcal{A}$ then $H^A \to H^B$ is an epimorphism in $\mathcal{L}(\mathcal{A}, \mathcal{B})$. Letting Q be a cogenerator for $\mathcal{L}(\mathcal{A}, \mathcal{B})$, it then suffices to show that $[H^B, Q] \to [H^A, Q]$ is a monomorphism in $\mathcal{B}$. But by 4.1 this is equivalent to $Q(B) \to Q(A)$ which is a monomorphism since Q is in $\mathcal{L}(\mathcal{A}, \mathcal{B})$.

THEOREM 4.4 (*The Full Imbedding Theorem*). *Every small abelian category* $\mathcal{A}$ *admits a full, exact, covariant imbedding into a category* $\mathcal{B}^{\mathcal{R}}$ *for some ring* $\mathcal{R}$.

Proof. Composing the functor $H: \mathcal{A} \to \mathcal{L}(\mathcal{A}, \mathcal{B})$ with the duality functor on the range category we obtain a full, exact, covariant imbedding H^* from $\mathcal{A}$ to a category with a projective generator. Now since $\mathcal{A}$ is small, by replacing the generator with the coproduct of a sufficient number of copies of itself we may assume that every object in the image of H^* is finitely generated. The theorem then follows by composing H^* with the functor T of 3.1.

5. Injective functors are not necessarily cokernel preserving. Let Q be injective in the category $(\mathcal{A},\mathcal{G})$ where $\mathcal{A}$ is a small abelian category, and consider an exact sequence in

$$A \to B \to C \to 0.$$

Then the sequence $0 \to H^C \to H^B \to H^A$ is exact in $(\mathcal{A},\mathcal{G})$ and so since Q is injective the sequence

$$[H^A, Q] \to [H^B, Q] \to [H^C, Q] \to 0$$

is exact in $\mathcal{G}$. But by 4.1 the last sequence is equivalent to $Q(A) \to Q(B)$ $Q(C) \to 0$. In other words an injective in $(\mathcal{A},\mathcal{G})$ is cokernel preserving as a functor. This fact is put to strong use in the work of Freyd [3], and in particular in the proof of 4.2. One may hope to generalize a number of the results in [3] by showing that if $\mathcal{B}$ is a reasonably nice abelian category, then injectives in $(\mathcal{A},\mathcal{B})$ are necessarily cokernel preserving. However, we are going to show that this need not be true even when $\mathcal{B}$ is a category of modules. In fact we shall show that an injective functor need not even preserve epimorphisms.

Let $\mathcal{A}$ and $\mathcal{A}'$ be any small additive categories, and consider the functor category $(\mathcal{A},(\mathcal{A}',\mathcal{G}))$. This last is the same as the category of covariant additive group-valued functors of two variables. For $A \in \mathcal{A}$ and $A' \in \mathcal{A}'$ define $H^A \otimes H^{A'} \in (\mathcal{A},(\mathcal{A}',\mathcal{G}))$ by $H^A \otimes H^{A'}(B,B') = [A,B] \otimes [A',B']$ where $B \in \mathcal{A}$, $B' \in \mathcal{A}'$, and the tensor product is over the integers. Then for $T \in (\mathcal{A},(\mathcal{A}',\mathcal{G}))$ it is not difficult to exhibit an isomorphism

$$[H^A \otimes H^{A'}, T] \approx T(A,A')$$

which is natural in A, A', and T (Cf. 4.1). Then precisely as in § 4 we can show that $\bigoplus_{(A,A') \in \mathcal{A} \times \mathcal{A}'} H^A \otimes H^{A'}$ is a projective generator for $(\mathcal{A},(\mathcal{A}',\mathcal{G}))$. Hence the latter being an A.B.5 category with a generator has an injective cogenerator. (By considering n-fold tensor products we can generalize this result to additive functors in any number of variables).

In particular let $\mathcal{A}'$ be a ring $\mathcal{R}$ (that is, an additive category with a single object A') and denote $H^A \otimes H^{A'}$ by $H^A \otimes \mathcal{R}$. Then the functor category $(\mathcal{A}',\mathcal{G})$ in this case is just the module category ${}^{\mathcal{R}}\mathcal{G}$. Also let $\mathcal{A}$ be a full subcategory $\mathcal{G}_0$ of $\mathcal{G}$ consisting of one member from every class of isomorphic finitely generated abelian groups. Then $\mathcal{G}_0$ is small, and it follows from the fact that $\boldsymbol{Z}$ is noetherian that $\mathcal{G}_0$ is abelian. Let Q be an injective cogenerator for the category $(\mathcal{G}_0, {}^{\boldsymbol{Z}_n}\mathcal{G})$ where $\boldsymbol{Z}_n$ denotes the ring of integers reduced modulo n. Consider the epimorphism $f\colon \boldsymbol{Z}_4 \to \boldsymbol{Z}_2$ in $\mathcal{G}_0$ defined by $f(1) = 1$. Then f induces a natural transformation

$$\eta\colon H^{\boldsymbol{Z}_2} \otimes \boldsymbol{Z}_2 \to H^{\boldsymbol{Z}_4} \otimes \boldsymbol{Z}_2.$$

Evaluating η at $\boldsymbol{Z}_4 \in \mathcal{G}_0$ we have

$$[\boldsymbol{Z}_2, \boldsymbol{Z}_4] \otimes \boldsymbol{Z}_2 \to [\boldsymbol{Z}_4, \boldsymbol{Z}_4] \otimes \boldsymbol{Z}_2 \tag{1}$$

Using the identifications $[\boldsymbol{Z}_2, \boldsymbol{Z}_4] \approx \boldsymbol{Z}_2$ and $[\boldsymbol{Z}_4, \boldsymbol{Z}_4] \approx \boldsymbol{Z}_4$, the morphism (1) becomes $g \otimes 1\colon \boldsymbol{Z}_2 \otimes \boldsymbol{Z}_2 \to \boldsymbol{Z}_4 \otimes \boldsymbol{Z}_2$ where g is multiplication by 2. It follows that $g \otimes 1 = 0$, whereas $\boldsymbol{Z}_2 \otimes \boldsymbol{Z}_2 \approx \boldsymbol{Z}_2$ is not zero. Hence $g \otimes 1$ is not a monomorphism. Since a monomorphism of functors must be pointwise a monomorphism it follows that η is not a monomorphism. Since Q is a cogenerator we then see that

$$[H^{\boldsymbol{Z}_4} \otimes \boldsymbol{Z}_2, Q] \to [H^{\boldsymbol{Z}_2} \otimes \boldsymbol{Z}_2, Q] \tag{2}$$

is not an epimorphism. But (2) is equivalent to $Q(\boldsymbol{Z}_4) \to Q(\boldsymbol{Z}_2)$. Therefore Q does not preserve epimorphisms.

However, if the ring $\mathfrak{R}$ is flat considered as an abelian group (that is, if $\mathfrak{R}$ is a torsion free group) then arguments similar to the above show that for any small abelian category an injective object in $(\mathcal{A}, {}^{\mathfrak{R}}\mathcal{G})$ is necessarily cokernel preserving.

REFERENCES.

[1] H. Cartan and S. Eilenberg, *Homological Algebra,* Princeton University Press (1956).

[2] B. Eckmann and A. Schopf, "Über injektive Moduln," *Archiv der Mathematik,* vol. 4 (1953), pp. 75-78.

[3] P. Freyd, ***Abelian Categories*** (Mimeographed Notes), Columbia University (1962).

[4] P. Gabriel, "Des Catégories Abéliennes," *Bulletin de la Société Mathématique de France,* vol. 90 (1962), pp. 323-448.

[5] A. Grothendieck, "Sur Quelques Points d'Algèbre Homologique," *Tôhoku Mathematical Journal,* vol. 9 (1957), pp. 119-221.

[6] S. Lubkin, "Imbedding of Abelian categories," *Transactions of the American Mathematical Society,* vol. 97 (1960), pp. 410-417.

ALMOST ACYCLIC MAPS OF MANIFOLDS.[1]

By Kyung Whan Kwun and Frank Raymond.[2]

1. Introduction. A map is called *proper* if the inverse image of each compact set is compact. A proper map is said to be *acyclic* over a coefficient domain L if each point inverse is cohomologically trivial over L. If a statement concerning acyclic maps does not mention L, L should be assumed to be an arbitrary but fixed principal ideal domain. The cohomology theory we use is the sheaf-theoretic one [3], and therefore, it may be interpreted as Cech or Alexander-Spanier cohomology if the space is paracompact or if it is locally compact and compact supports are used. We use the homology theory defined by Borel and Moore in [2]. To avoid repetition, we will say that a map $f\colon (X,A) \to (Y,B)$ is *admissible* if (1) f is proper, (2) (X,A), (Y,B) are locally compact Hausdorff closed pairs and (3) $f(A) = B$ and $f(X-A) = Y-B$.

The first part of the paper deals with the following question:

Suppose $f\colon (X,A) \to (Y,B)$ is an admissible map such that $f \mid X-A$ is acyclic. Under what conditions is f acyclic?

Since even when f is a relative homeomorphism, anything could happen to A, the above question is considered by imposing some structure on X and demanding f to preserve it.

The two theorems along this line are the following:

Theorem 1. *Let $f\colon (X,A) \to (Y,B)$ be an admissible map, where A and B are collared in X and Y, respectively. If $f \mid X-A$ is acyclic then f is acyclic.*

Theorem 2. *Let X, Y be locally orientable generalized n-manifolds with boundary and $f\colon (X, \operatorname{Bd} X) \to (Y, \operatorname{Bd} Y)$ be an admissible map. If $f \mid \operatorname{Int} X$ is acyclic then f is acyclic.*

Recall there exists an example of a proper map of euclidean 3-space,

---

Received February 3, 1964.

[1] Part of this paper was presented to the American Mathematical Society at Boulder, Colorado, August 29, 1963, under the title of "Sufficient conditions that maps be acyclic."

[2] Partially supported by NSF GP-626, NSF GP-812 and U. S. Air Force Office of Scientific Research.

E^3, onto itself which is not acyclic even though it is a homeomorphism between two dense open subsets of E, see [1]. Theorem 1 or 2 implies that in a similar situation in case of manifolds with boundary, "wrong" point inverses, if they exist, cannot be confined to the boundary only.

Most of the remainder of the paper studies the cohomology of the set A when $f\colon (X, A) \to (Y, B)$ is an admissible map between manifolds.

The next theorem imposes strong cohomological conditions on upper semicontinuous decompositions (proper maps) of manifolds for which the decomposition spaces are again manifolds.

THEOREM 3. *Let X be a connected orientable generalized n-manifold and Y a locally orientable generalized n-manifold ($n - gm$). Suppose $f\colon (X, A) \to (Y, B)$ is an admissible map such that $f \mid X - A$ is acyclic, and $A \neq X$. Then*

(1) *$H_c^p(Y)$ is a direct summand of $H_c^p(X)$ for each p;*

(2) *if $H_c^q(B) = 0$, $q > 0$, then the sequence*

$$0 \to H_c^p(Y) \xrightarrow{f^*} H_c^p(X) \xrightarrow{i^*} H_c^p(A) \to 0$$

is exact and splits for $p > 0$ and

$$f^*\colon H_c^0(B) \to H_c^0(A)$$

is an isomorphism;

(3) *if $H_c^q(X) = 0$, $q \neq 0, n$, then*

$$f^*\colon H_c^p(B) \to H_c^p(A)$$

is an isomorphism for all p.

We have the following corollaries to Theorem 3.

COROLLARY 1. *Let X be a compact connected orientable $n - gm$ and A be a closed subset of X. Let Y be the space obtained from X by shrinking A to a point. If Y is a gm then, $H_c^*(X) \not\approx H_c^*(Y)$ if and only if A is cohomologically non-trivial.*

This answers in the negative the question whether one can shrink to a point a cohomologically non-trivial continuum in some compact manifold and still get a homeomorphic space. John Hempel points out that this question is answered in the affirmative if the compactness is dropped.

COROLLARY 2. *Let $f\colon X \to Y$ be a proper map of an orientable $n - gm$ X onto an $n - gm$ Y. Further suppose that $H_c^*(X)$ is isomorphic to either $H_c^*(S^n)$ or $H_c^*(E^n)$. Then the singular set B of f cannot be 0-dimensional.*

Here the singular set of f is the set $Cl\{y \in Y \mid f^{-1}(y)$ is not cohomologically trivial$\}$ and the 0-dimensionality is used in the sense that each point has an arbitrarily small neighborhood which is both open and closed.

COROLLARY 3. *Let $f\colon X \to Y$ be a proper map of an orientable $n - gm$ X onto a $n - gm$ Y. Suppose either $H_c^*(X) = H_c^*(S^n)$ or $H_c^*(E^n)$. If $H_c^*(Y) \not\approx H_c^*(X)$ then the singular set of f is all of Y.*

Thus if one starts with an $n - gm$ like E^n or S^n, one cannot change the homology without having *every* point of X belonging to the closure of the union of the cohomologically non-trivial inverse sets.

COROLLARY 4. *Let X be a generalized n-cell and A be a closed subset of X. If the space Y obtained from X by shrinking A to a point is a locally orientable generalized manifold with non-empty boundary then A is cohomologically trivial.*

A manifold may be mapped onto a manifold of higher dimension. Theorem 4, whose proof is similar to the proof of Theorem 3, restricts the possibility of such phenomena.

THEOREM 4. *Let $f\colon X \to Y$ be a proper map of a connected orientable $n - gm$ onto an $m - gm$, $m > 0$. Let $\tilde{H}^k(f^{-1}(y)) = 0$, $k \leq s$, $n \leq 2(s+1)$ all $y \in Y$. Then .*

(a) *either f is acylic and $n = m$, or else $m \leq s$ when $n < 2(s+1)$ and $m \leq s+1$ when $n = 2(s+1)$;*

(b) *if there exists $y_0 \in Y$ such that $H_c^{s+1}(f^{-1}(y_0)) \neq 0$ (and if $n = 2(s+1)$ with $m \leq s$, then also assume that $H_{s+1}{}^c(X) = 0$), then $n = m + (s+1)$.*

Theorem 3 leads one to suspect that, in general, the inverse image A of the singular set must satisfy a homological condition resembling Poincaré duality. To be precise let $f\colon X \to Y$ *satisfy the hypothesis* (2) *of Theorem* 3 *and let X be compact.* Then

THEOREM 5. *The diagram* $(p \neq 0, n)$

$$\begin{tikzcd}
0 \arrow[r] & H^p(Y) \arrow[r, "f^*"] & H^p(X) \arrow[r, "i^*"] & H^p(A) \arrow[r] & 0 \\
0 & H_{n-p}(Y) \arrow[l] \arrow[u, "\Delta"] & H_{n-p}(X) \arrow[l, "f_*"'] \arrow[u, "\Delta"] & H_{n-p}(A) \arrow[l, "j_*"'] \arrow[u, "\psi"] & 0 \arrow[l]
\end{tikzcd} \tag{5.1}$$

is commutative, the rows split and the vertical maps are isomorphisms induced by Poincaré duality. In particular, all modules are finitely generated.

2. Proof of Theorem 1. If $f: X \to Y$ is a proper acyclic map of a locally compact space X onto a locally compact space Y, then $f^*: H^*(Y) \to H^*(X)$ is an isomorphism. For paracompact X and Y this is well known and we suspect that the general case is also known. A simple proof is given in [5]. We also recall that A is collared in X if there exists homeomorphism h of $A \times [0,1)$ onto an open subset of X such that $h(a,0) = a$ for each $a \in A$.

To prove Theorem 1, it suffices to show that for each $b \in B$, $f^*: H^*(b) \to H^*(f^{-1}(b))$ is an epimorphism.

For any fundamental system $\{U_\alpha\}$ of neighborhoods of b, $\{V_\alpha\}$, $V_\alpha = f^{-1}(U_\alpha)$ is a fundamental system of neighborhoods of $f^{-1}(b)$. And therefore,

$$\begin{aligned} H^*(b) &= \operatorname{dir\,lim} H^*(U_\alpha) \\ H^*(f^{-1}(b)) &= \operatorname{dir\,lim} H^*(V_\alpha). \end{aligned} \tag{1}$$

Let $c \in H^*(f^{-1}(b))$. Then c is represented, in the relation (1), by some $c' \in H^*(V_\alpha)$. Let V be a neighborhood of $f^{-1}(b)$ such that $V \subset V_\alpha$ and $V = (V \cap A) \times [0,1)$ with each $v \in V \cap A$ corresponding to $(v,0)$. We may also assume, with no loss of generality, that $U_\alpha = (U_\alpha \cap B) \times [0,1)$ with each $u \in U_\alpha \cap B$ corresponding to $(u,0)$.

Consider the following commutative diagram.

$$\begin{tikzcd}
H^*(U_\alpha) \arrow[r, "i_\alpha'"] \arrow[d, "f_\alpha^*"] & H^*(U_\alpha - B) \arrow[d, "\bar{f}_\alpha^*"] \\
H^*(V_\alpha) \arrow[r, "i_\alpha"] \arrow[d, "i'"] & H^*(V_\alpha - A) \arrow[d, "i''"] \\
H^*(V) \arrow[r, "i"] & H^*(V - A)
\end{tikzcd}$$

where all homomorphisms are induced by f or inclusion, and i_α', $\bar{f}_\alpha^*$ and i are isomorphisms.

Now c is also represented by $c'' = i'(c')$. There exists an element $d_1 \in H^*(U_\alpha)$ such that

$$i_\alpha f_\alpha^*(d_1) = i_\alpha(c').$$

Hence $i' f_\alpha^*(d_1) = c''$. Therefore $f_\alpha^*(d_1) \in H^*(V_\alpha)$ represents $c \in H^*(f^{-1}(b))$. Thus c is the image under the homomorphism $H^*(b) \to H^*(f^{-1}(b))$ of the element of $H^*(b)$ represented by $d_1 \in H^*(U_\alpha)$.

3. Proof of Theorem 2. For each $b \in \operatorname{Bd} Y$, there exists an arbitrarily small connected neighborhood U such that $U - \operatorname{Bd} Y$, $U \cap \operatorname{Bd} Y$ are connected orientable *gms* [6, Lemma 2.2]. For such a U, take $V = f^{-1}(U)$.

Consider the following commutative diagram

$$\begin{tikzcd}
H^p(U) \arrow[r, "i"] \arrow[d, "f^*"] & H^p(U - \operatorname{Bd} Y) \arrow[d, "f'^*"] \\
H^p(V) \arrow[r, "j"] & H^p(V - \operatorname{Bd} X)
\end{tikzcd}$$

By [8, Theorem 6, p. 40] i and j are isomorphisms. Since f'^* is an isomorphism [5], f^* also is. By passing to the direct limit we obtain

$$H^p(f^{-1}(b)) \cong H^p(b)$$

4. Proof of Theorem 3. (1) If $X \neq A$, one can show that

$$f_*\colon H_n(X) \to H_n(Y)$$

is an isomorphism. In fact, (1) of Theorem 3 is true if f_* is an isomorphism as just stated.

As in Théorème 7.9, p. 510 of [3], let $\xi = \Delta^{-1}(1)$, $1 \in H^*(X)$, and ξ' be the one corresponding to Y. Then $f_*(\xi) = \pm \xi'$. In the diagram

(i)

$$\begin{tikzcd}
H_c^p(Y) \arrow[r, "\Delta'^{-1}"] \arrow[d, "f^*"] & H_{n-p}^c(Y) \\
H_c^p(X) \arrow[r, "\Delta^{-1}"] & H_{n-p}^c(X) \arrow[u, "f_*"']
\end{tikzcd}$$

$\Delta' f_* \Delta^{-1} f^* = \pm$ identity, where Δ^{-1}, Δ'^{-1} are inverses of the Poincaré duality isomorphisms. This fact follows from the fact that Δ^{-1}, (Δ'^{-1}), are obtained by capping (see [3]) with ξ, (ξ'), and $f_*(\xi) = \pm \xi'$. Thus f^* in (i) is a splitting monomorphism and f_* is a splitting epimorphism.

(2) Consider the following commutative diagram

(ii)

$$\begin{tikzcd}
{} \arrow[r] & H_c^{p-1}(A) \arrow[r, "d^*"] & H_c^p(X - A) \arrow[r, "j^*"] & H_c^p(X) \arrow[r, "i^*"] & H_c^p(A) \arrow[r, "d^*"] & H_c^{p+1}(X - A) \arrow[r] & {} \\
{} \arrow[r] & H_c^{p-1}(B) \arrow[r, "d'^*"] \arrow[u, "f''^*"] & H_c^p(Y - B) \arrow[r, "j'^*"] \arrow[u, "f'^*"] & H_c^p(Y) \arrow[r, "i'^*"] \arrow[u, "f^*"] & H_c^p(B) \arrow[r, "d'^*"] \arrow[u, "f''^*"] & H_c^{p+1}(Y - B) \arrow[r] \arrow[u, "f'^*"] & {}
\end{tikzcd}$$

where the rows are exact and $f'' = f \mid A$, $f' = f \mid X - A$. We have just seen that f^* is a splitting monomorphism. Furthermore, f'^* is an isomorphism. Hence, if $H_c^p(B) = 0$, then d^* is trivial in dimension p. Thus for $p > 1$, the sequence

$$0 \to H_c^p(X - A) \xrightarrow{j^*} H_c^p(X) \xrightarrow{i^*} H_c^p(A) \to 0$$

is exact, and j'^* is an isomorphism. Since $j^* f'^* j'^{*-1} = f^*$, the sequence

$$0 \to H_c^p(Y) \xrightarrow{f^*} H_c^p(X) \xrightarrow{i^*} H_c^p(A) \to 0$$

splits, $p > 1$.

For $p = 1$, i'^* and d^* are trivial. We claim kernel $j^* = f'^*(\text{kernel}\, j'^*)$. This follows because f^* is a monomorphism and f'^* an isomorphism. We can split the diagram (ii) into the two diagrams

$$\begin{tikzcd}
0 \arrow[r] & H_c^1(X - A)/\text{kernel}\, j^* \arrow[r] & H_c^1(X) \arrow[r, "i^*"] & H_c'(A) \arrow[r] & 0 \\
0 \arrow[r] & H_c^1(Y - B)/\text{kernel}\, j'^* \arrow[r] \arrow[u, "f_1'^*"'] & H_c^1(Y) \arrow[r] \arrow[u, "f^*"'] & 0 \arrow[u, "f''^*"'] &
\end{tikzcd}$$

and

$$\begin{tikzcd}
0 \arrow[r] & H_c^0(X) \arrow[r] & H_c^0(A) \arrow[r, "d^*"] & \text{kernel}\, j^* \arrow[r] & 0 \\
0 \arrow[r] & H_c^0(Y) \arrow[r] \arrow[u, "f^*"'] & H_c^0(B) \arrow[r] \arrow[u, "f''^*"'] & \text{kernel}\, j'^* \arrow[r] \arrow[u, "f'^*"'] & 0.
\end{tikzcd}$$

The homomorphism $f_1'^*$ is induced by f'^* and is an isomorphism. Since the rows are exact, it is easily seen that $\operatorname{Im} i^*$ *is* the cokernel of **f***. The five lemma implies that f''^* is an isomorphism in the second of the two diagrams.

(3) Consider the diagram (ii) once again. $H_c^p(X) = H_c^p(Y) = 0$, $p \neq 0, n$. The five lemma then implies that $f''^* \colon H_c^p(B) \to H_c^p(A)$ is an isomorphism for all p.

Remark. The homology analogue of (2) is the splitting of the exact sequence

$$0 \to H_{n-p}{}^c(Y) \xrightarrow{j_* f_*'^{-1} j_*'^{-1}} H_{n-p}{}^c(X) \xrightarrow{i_*} H_{n-p}{}^c(X, X - A) \to 0, \text{ for } p \neq 0.$$

Here $f_*(j_* f_*'^{-1} j_*^{-1})$ is the identity and the homomorphisms are just the duals to those in cohomology. One can prove this by replacing diagram (ii) by its homology analogue and copying the argument with a slight difficulty for $p=1$ which is overcome by a similar argument as used for cohomology at $p=1$. However, much more satisfying is to consider both diagrams and the isomorphisms between them induced by Poincaré duality [8; 6.2]. What must be shown is that this three dimensional diagram commutes (or anti-commutes). This is done by showing that the isomorphisms are induced by cap products.

If instead of compact supports we used closed supports and assumed X and Y paracompact then we would, by the same argument, be able to prove an analogous theorem in terms of closed supports.

We leave to the reader the details of the proofs of Corollaries 1-4. We mention, however, that in the proof of Corollary 1 one observes that $H_c^*(X) \approx H^*(X)$ is finitely generated for each dimension and that any finitely generated module over a principal ideal domain cannot be isomorphic to a proper factor module.

5. Proof of Theorem 4. (a) First suppose $m > s$. We have

$$H_p^c(Y, Y-y) = 0, \text{ for all } p \leqq s.$$

The map $f_*'': H_p^c(X, X-f^{-1}(y)) \to H_p^c(Y, Y-y)$ is an isomorphism, for all $p \leqq s$ and is onto in dimension $s+1$. Poincaré duality [8; 6.1] implies that $H_c^k(f^{-1}(y)) = 0$, for $k = n, n-1, \cdots, n-s$. But, if $n-s \leqq s+1$, f is acyclic, since acyclic maps cannot raise dimension, we are forced to conclude that either $m=n$ and f is acyclic or else f is not acyclic and $m \leqq s$. In case we only had assumed $n-s \leqq s+2$, we could have trouble when $n = 2s+2$ in dimension $s+1$. However it isn't hard to check then that either f is acyclic or $m = s+1 < n$.

(b) We saw in (a) that if f is not acyclic then $m \leqq s$ if $n < 2s+2$, and $m \leqq s+1$ if $n = 2s+2$. If $m \leqq s$, then $H_k^c(Y, Y-y) \approx H^{n-k}(f^{-1}(y)) = 0$, for $k < m$, and $H^{n-m}(f^{-1}(y)) = L$. On the other hand, $H_{n-(s+1)}^c(Y, Y-y_0) \approx H^{s+1}(f^{-1}(y_0))$, if $n-(s+1) \leqq s$, or equivalently if $n \leqq 2s+1$. Therefore we have shown that $m+(s+1) = n$. It remains to check the case $n = 2s+2$ and $m \leqq s$. However, if $H_{s+1}^c(X) = 0$, we get a contradiction implying that no such f can exist in this case.

Notice that the Leray sheaf of the map f is 0 in all dimensions except $s+1$ and there it is constant isomorphic to L. Nevertheless examples easily show that $f^{-1}(y)$ need not be generalized manifolds.

In dimensions $p=1$, 2, 3, and 4, the bounds by s give the best possible results for mappings $S^p \to S^{p+k}$, $k > 0$.

6. Proof of Theorem 5. In diagram 5.1 of §1 we have from (2) of Theorem 3 that the top row splits. From the remarks following the proof of Theorem 3 we have (using closed supports) that the bottom row of (5.1) splits Furthermore the first rectangle defines the splitting and is cummutative. Obviously there is only one way to define ψ so that the diagram is commutative. Furthermore ψ preserves the splitting properties and must be an isomorphism.

Remark. For a compact set A one always has $H^p(A) = \sum_{A' \subset A} H^p(A')$, $H_{n-p}(A) = \prod_{A' \subset A} H_{n-p}(A')$, where A' denotes a connected component of A and the sum and product runs over the components of A. In Theorem 5, $H^p(A)$ and $H_{n-p}(A)$ are finitely generated, and therefore both representations are *direct sums.* Furthermore, f maps the components A' of A onto the components of B in a 1:1-fashion. Therefore the isomorphism ψ induced by f splits naturally into the direct product (sum) of the isomorphisms $\psi_{A'}: H_{n-p}(A') \to H^p(A')$.

7. Manifolds which are smash products. In this section, we include a heretofore unpublished result of F. Raymond which serves as an illustration of Theorems 3 and 5. In view of its independent interest, we give a proof.

The smash product (or reduced join) $X \wedge Y$ of spaces X and Y is the product $X \times Y$ modulo $x \times Y \cup X \times y$, $x \in X$, $y \in Y$. Strictly speaking, therefore, X, Y, should be *based* spaces. However the choice of base points is irrelevant to our theorem and therefore will be ignored. We will assume Y is a connected $n-gm$ and $Y = A \wedge B$, where $\dim A$, $\dim B > 0$. Then the compactness of A and B follows from the local compactness of Y.

THEOREM 6. *Let $Y = A \wedge B$ be a connected paracompact $n-gm$, $(n > 0)$, and homology locally connected in the sense of singular homology* (hlc). *Then A is a $k-gm$, B is an $(n-k)-gm$ and A, B and Y are all spherelike.*

COROLLARY. *If Y, A, and B are locally euclidean, they are all homotopy spheres.*

Proof. Denote by y the image point of $A \times b \cup a \times B$, $a \in A$, $b \in B$, under the decomposition map $f: A \times B \to A \wedge B$. The map $f \mid (A - a) \times (B - b)$ is a homeomorphism and so

$$(U - a) \times (B - b) \cup (A - a) \times (V - b)$$

is a deleted neighborhood of y in Y, where U is a neighborhood of a in A and V a neighborhood of b in B. Since Y is locally orientable, y is contained in some orientable neighborhood. Using [6; Theorem 6] it follows that $(A-a)$, $(B-b)$, is a connected orientable $k-gm$, $((n-k)-gm)$. Since $Y-y$ is orientable Y must be orientable. Furthermore, since Y is compact, the end cohomology groups $I^p(Y-y)\approx H^p(S^{n-1})$ by [7]. From [7;2.2] it follows that the identity map induces an isomorphism

$$i\colon H_c{}^p(Y-y)\rightarrow H^p(Y),\, p\neq 0, n.$$

In order to establish the theorem for an arbitrary principal ideal domain it will suffice to prove the theorem for an arbitrary field of coefficients as may be justified by a forthcoming paper of one of the authors. Assume L is a field. It follows that $Y-y$ has finitely generated cohomology with compact or closed supports. (See [7; p. 959]). This implies that both factors $(A-a)$ and $(B-b)$ of $(Y-y)$ have finitely generated cohomology. We wish to use the Künneth theorem for both compact and closed supports. The Künneth theorem is valid for compact supports and is valid for ordinary singular cohomology when the cohomology is finitely generated. But Y is *hlc* and so both $A-a$ and $B-b$ are also *hlc*. The sheaf theoretic and singular cohomologies agree for any paracompactifying family. Y is actually metrizable by the paracompact and *hlc* assumption and so in particular $Y-y$ is paracompact. We write the isomorphisms i as the isomorphisms

$$i_A\otimes i_B\colon \sum_{s+t=p} H_c{}^s(A-a)\otimes H_c{}^t(B-b)\rightarrow \sum_{s+t=p} H^s(A-a)\otimes H^t(B-b)$$

where $p\neq 0, n$. The naturality of this isomorphism forces the $k-gm$ $A-a$ and the $(n-k)-gm$ $B-b$ to have trivial compact cohomology except in the top dimensions, (the homomorphism $H_c{}^k(A-a)\rightarrow H^k(A-a)$ is trivial).

In particular, $A-a$, $B-b$ and therefore $Y-y$ all must have trivial cohomology with closed supports. By [7; 4.7], A and B and Y are all spherelike *gms*.

Remarks. Conversely, if A and B are spherelike *gms* then $Y-y = (A-a)\times(B-b)$ is an acyclic *gm*. By [7], Y, the one point compactification of $Y-y$, must be a *gm* (and in fact sphere-like).

In the theorem if A, B or Y is of dimension $\leqq 2$, then it must be an actual sphere of that dimension. Now assume that A and B have the same homotopy type as a $C-W$ complex (true if A and B are ANR's) and $L=Z$, the integers. Then Y is simply connected (see for example [9; 3.16]) and therefore has the homotopy type of an actual sphere. Furthermore, one can show that A and B also have the same homotopy type as spheres if one assumes, when dimensions are $\geqq 3$, that the local fundamental groups at the base points are trivial. Hence, in conclusion we can say that the classical construction

$$S^k \times S^{n-k} \xrightarrow{f} S^k \wedge S^{n-k} = S^n$$

uniquely characterizes, modulo the Poincaré conjecture, the construction

$$A \times B \xrightarrow{f} A \wedge B = Y$$

if A, B and Y are all locally euclidean.

Also, it is not difficult to give examples where Y is a sphere but A and B fail to be locally euclidean.

8. Remarks on factors of cubes. We like to take this opportunity to make some remarks on our earlier work: "Factors of cubes," *American Journal of Mathematics*, vol. 84 (1962), pp. 433-440, which will be hereafter referred to as FC.

Question 1 of [FC; p. 437] has been neatly solved in the affirmative by Jehpill Kim of Seoul National University, in a paper to appear. It is shown in [FC] that for each $n \geqq 8$, there exist spaces A, B which are not manifolds with boundary such that $A \times B = I^n$. We have conjectured [FC, the last sentence of Section 2, p. 437] that it is possible for $n = 6$ and 7. The following theorem, which incidentally supersedes Theorem 2 of [FC], proves this conjecture and also answers the second half of Question 2 [FC, p. 437] in the affirmative:

THEOREM 7. *Let $\alpha \subset \operatorname{Int} I^m$, $\beta \subset \operatorname{Int} I^n$ be arcs and I^m/α, I^n/β respectively denote the spaces obtained from I^m and I^n by shrinking α and β to a point. Then $I^m/\alpha \times I^n/\beta = I^{m+n}$, for all $m, n \geqq 2$.*

We first prove a lemma in which the following notation will be used. Let $f\colon I^m \to I^m/\alpha$ be the natural projection. For any homeomorphism $h\colon S^n \to S^n$, let $f \times h\colon I^m \times S^n \to I^m/\alpha \times S^n$ be the map such that $(f \times h)(a, b) = (f(a), h(b))$.

LEMMA. *For any positive integer n, and homeomorphism $h\colon S^n \to S^n$, there exists a homeomorphism g of $I^m \times S^n$ onto $I^m/\alpha \times S^n$ such that $g \mid (\operatorname{Bd} I^m \times S^n) = (f \mid \operatorname{Bd} I^m) \times h$.*

Proof. We use induction on n. For $n = 1$, it is essentially the same as (2) of the proof of Theorem 2 of [FC]. Precisely, there exists a continuous map

$$\mu \colon I^m \times S^1 \xrightarrow{\text{onto}} I^m \times S^1$$

such that $\mu \mid (\operatorname{Bd} I^m \times S^1) = 1$ and the point inverses under μ are precisely those of

$$f \times 1 \colon I^m \times S^1 \to I^m/\alpha \times S^1.$$

Then $g' = (f \times 1)\mu^{-1}$ gives a homeomorphism of $I^m \times S^1$ onto $I^m/\alpha \times S^1$ such that $g' \mid (\operatorname{Bd} I^m \times S^1) = f \times 1$. Let $g = (1 \times h)g'$.

We now suppose the lemma is true for $n - 1 \geqq 1$, and prove it for n. Let $S^n = D_1 \cup D_2$, where D_1 and D_2 are n-cubes such that $\operatorname{Bd} D_1 = \operatorname{Bd} D_2$. There exists a homeomorphism

$$g' \colon I^m \times \operatorname{Bd} D_1 \xrightarrow{\text{onto}} I^m/\alpha \times \operatorname{Bd} D_1.$$

such that

$$g' \mid (\operatorname{Bd} I^m \times \operatorname{Bd} D_1) = (f \mid \operatorname{Bd} I^m) \times 1$$

Consider the homeomorphisms

$$f_i = f \mid \operatorname{Bd} I^m \times 1 \colon \operatorname{Bd} I^m \times D_i \xrightarrow{\text{onto}} \operatorname{Bd}(I^m/\alpha) \times D_i.$$
$$f_i = g' \text{ on } \operatorname{Bd} I^m \times \operatorname{Bd} D_1 = \operatorname{Bd} I^m \times \operatorname{Bd} D_2.$$

Hence there exist homeomorphisms

$$g_i \colon \operatorname{Bd}(I^m \times D_i) \xrightarrow{\text{onto}} \operatorname{Bd}(I^m/\alpha \times D_i)$$

such that

$$g_i \mid (\operatorname{Bd} I^m \times D_i) = f_i$$

Now, by Theorem 2 of [FC], as $n \geqq 2$, $I^m/\alpha \times D_i$ and $I^m/\alpha \times D_i$ are $(m+n)$-cells. Hence g_i can be extended to homeomorphisms

$$g_i' \colon I^m \times D_i \to I^m/\alpha \times D_i$$

Since $g_i' \mid (I^m \times \operatorname{Bd} D_i)$ are both g', they give rise to a homeomorphism

$$g'' \colon I^m \times S^n \xrightarrow{\text{onto}} I^m/\alpha \times S^n$$

such that $g'' \mid \operatorname{Bd} I^m \times S^n = (f \mid \operatorname{Bd} I^m \times S^n) \times 1$. Set $g = (1 \times h)g''$ to obtain the desired homeomorphism.

To prove the theorem let

$$f_1\colon I^m \to I^m/\alpha$$
$$f_2\colon I^n \to I^n/\beta$$

be the projections and

$$f_1'\colon \operatorname{Bd} I^m \to \operatorname{Bd}(I^m/\alpha)$$
$$f_2'\colon \operatorname{Bd} I^n \to \operatorname{Bd}(I^n/\beta)$$

be their restrictions. We will prove that

(1) $\operatorname{Int}(I^m/\alpha \times I^n/\beta) = E^{m+n}$

(2), $I^m/\alpha \times I^n/\beta$ is a manifold with boundary and

(3) $\operatorname{Bd}(I^m/\alpha \times I^n/\beta) = S^{m+n-1}$.

Since (1) and (2) imply that $I^m/\alpha \times I^n/\beta$ can be imbedded in euclidean $(m+n)$-space, E^{m+n}, our theorem follows as in [FC].

(1) is proved in a paper by one of the authors: "Product of Euclidean spaces modulo an arc," to appear in *Annals of Mathematics.*

(2) It suffices to check each point of $\operatorname{Bd}(I^m/\alpha \times I^n/\beta)$ has a "nice" neighborhood. Suppose $x \in \operatorname{Bd}(I^m/\alpha) \times I^n/\beta$. Then, using Theorem 2 of [FC], it has a neighborhood homeomorphic to $I^m \times I^n/\beta = I^{m+n}$. Similarly for $y \in I^m/\alpha \times \operatorname{Bd}(I^n/\beta)$.

(3) $\operatorname{Bd}(I^m/\alpha \times I^n/\beta) = \operatorname{Bd}(I^m/\alpha) \times I^n/\beta \cup I^n/\alpha \times \operatorname{Bd}(I^n/\beta)$.
By the lemma, there exists homeomorphisms

$$g_1\colon \operatorname{Bd} I^m \times I^n \xrightarrow{\text{onto}} \operatorname{Bd}(I^m/\alpha) \times I^n/\beta$$

such that $g_1 \mid \operatorname{Bd} I^m \times \operatorname{Bd} I^n = f_1' \times f_2'$ and

$$g_2\colon I^m \times \operatorname{Bd} I^n \xrightarrow{\text{onto}} I^m/\alpha \times \operatorname{Bd}(I^n/\beta)$$

such that $g_2 \mid \operatorname{Bd} I^m \times \operatorname{Bd} I^n = f_1' \times f_2'$. Hence g_1 and g_2 combine to give a homeomorphism of $\operatorname{Bd}(I^m \times I^n)$ onto $\operatorname{Bd}(I^m/\alpha \times I^n/\beta)$.

FLORIDA STATE UNIVERSITY
AND
UNIVERSITY OF MICHIGAN.

REFERENCES.

[1] R. H. Bing, "Decompositions of E^3," *Topology of 3-manifolds and Related Topics* (Editor: M. K. Fort, Jr.), Prentice Hall, Inc. (1962), pp. 5-21.

[2] A. Borel and J .C. Moore, "Homology theory for locally compact spaces," *Michigan Mathematical Journal*, vol. 7 (1960), pp. 249-256.

[3] Armand Borel and Andre Haefliger, "La classe d'homologie fondamentale d'un espace analytique," *Bulletin de la Société Mathématique de France*, vol. 89 (1961), pp. 461-513.

[4] Roger Godment, *Topologie algébrique et théorie des faisceaux*, Hermann, 1958, Paris.

[5] K. W. Kwun, *Remarks on injective sheaves* (unpublished).

[6] Frank Raymond, "Separation and union theorems for generalized manifolds with boundary," *Michigan Mathematical Journal*, vol. 7 (1960), pp. 7-21.

[7] ———, "The end point compactification of manifolds," *Pacific Mathematical Journal*, vol. 10 (1960), pp. 941-963.

[8] ———, "Local cohomology groups with closed supports," *Mathematische Zeitschrift*, vol. 76 (1961), pp. 31-41.

[9] George W. Whitehead, "Generalized homology theories," *Transactions of the American Mathematical Society*, vol. 102 (1962), pp. 227-283.

THE CLOSURE OF TRANSLATES IN l^p.

By D. J. Newman.[1]

1. Introduction. Let $f(x)$ have an absolutely convergent fourier series, and let $P(x)$ range over the set of trigonometric polynomials. What is the closure of the functions $P(x)f(x)$ in the l^p norm? In particular when is it al lof l^p?

This problem, the so-called 'closure of translates problem,' has been considered by various authors. Wiener (see [9]) treated the cases $p=1$ and $p=2$ and proved

(1.1) For $p=1$, this closure is all of l^1 if S, the set of zeros of $f(x)$, is empty.

(1.2) For $p=2$, this closure is all of l^2 if S has measure 0.

The converses of (1.1) and (1.2) are also true.

For other p, $1 \leqq p \leqq 2$, Beurling (see [1]) proved[2]

(1.3) The closure is all of l^p if S has Hausdorff dimension less than $2(p-1)/p$.

In the case of (1.3), however, the converse is not true. Indeed it is not difficult to produce a set S of Hausdorff dimension 1 which nevertheless has the aforementioned closure property for all $p>1$. (See our Corollary 2 to Theorem 4.)

In Section 2 of this paper we reinvestigate the case of $1 \leqq p \leqq 2$ and obtain some results which are in a sense both stronger and weaker than the Beurling theorem. Section 3 is concerned with the case $2<p<\infty$. (The surprising result in this connection is that there are sets, S, of *positive* measure for which, nevertheless, the closure is all of l^p). Finally, Section 4 treats the case $p=\infty$, properly modified, and the necessary and sufficient condition is obtained. (It is simply that S be nowhere dense).

Received April 24, 1963.

Revised April 24, 1964.

[1] The author was partially supported by the National Science Foundation NSF GP-02.

[2] Actually Beurling worked with L^p rather than l^p, but the distinction is not vital.

Definitions and basic notions. For $1 \leqq p < \infty$, l^p is defined to be the Banach space of sequences $\{c_n\}, -\infty < n < \infty$, normed by

$$\|\{c_n\}\|_p = (\sum_{-\infty}^{\infty} | c_n |^p)^{1/p}.$$

If $f(x)$ is a summable function (on $[0, 2\pi)$) we say that $f(x) \in l^p$ or $f(x)$ is an l^p function if its Fourier coefficients, $\{c_n\}$, are in l^p. In this case we understand $\| f(x) \|_p$ to mean $\| \{c_n\} \|_p$.

We say that $f(x)$ is Lip 1 if $f(x)$ has period 2π and if there exists an $M \ni | f(x) - f(y) | \leqq M | x - y |$ for all x, y. Note that for $P \geqq p$, $l^P \supseteq l^p$ and that, in particular, all the l^p spaces contain l^1.

The space l^1, when thought of as the space of functions with absolutely convergent fourier series, is actually a Banach algebra.

If I is any closed ideal in l^1, we denote by S_I the set of common zeros of the functions of I. We will make frequent use of the following:

LEMMA 1. *If* $f(x)$ *is* Lip 1 *and* $f(x) = 0$ *for all* $x \in S_I$ *then* $f(x) \in I$.

For proof see e.g. [6] or [7].

Now let S be a closed subset of $[0, 2\pi)$.

We call S a p-spanning-set if every closed ideal I in l^1 with $S_I = S$ is dense in l^p.

LEMMA 2. *S is a p-spanning-set if and only if there exist* Lip 1 *functions* $f_1(x), f_2(x), \cdots$ *each vanishing on S, such that*

$$\| f_n(x) - 1 \|_p \to 0.$$

Proof. If I is any closed ideal in l^1 with $S_I = S$ then, by Lemma 1, I contains all the functions $P(x)f_n(x)$, $P(x)$ any trigonometric polynomial. Since $\| P(x)f_n(x) - P(x) \|_p \to 0$ it follows that the l^p closure of I contains all the trigonometric polynomials and, since the polynomials are dense in l^p, so is I.

Conversely choose I as the l^1 closure of all Lip 1 functions which vanish on S. If S is a p-spanning-set then I is dense in l^p, and so the Lip 1 functions vanishing on S are dense in l^p. In particular the constant function, 1, is an l^p limit of such functions.

For the concepts of Hausdorff dimension and Hausdorff α-measure we refer the reader to [5].

The space l^∞ is defined as the Banach space of all bounded sequences, $\{c_n\}$, $-\infty < n < \infty$, normed by $\| \{c_n\} \|_\infty = \sup_{-\infty<n<\infty} | c_n |$.

We also introduce l_0^∞, defined as the closed subspace of l^∞ consisting of all sequences $\{c_n\}$ with $\lim_{n\to\pm\infty} c_n = 0$.

l_0^∞ can also be thought of as the closure of the polynomials under the l^∞ norm.

Finally, by a triangle function, we mean a function of the form

$$f(x) = \begin{cases} 0 \text{ outside } (\xi-\delta, \xi+\delta) \\ h(1-|x-\xi|/\delta) \text{ inside } (\xi-\delta, \xi+\delta) \end{cases}$$

where $0 \leqq \xi-\delta < \xi+\delta \leqq 2\pi$, $f(x)$ having period 2π. h is called the height and 2δ the base of $f(x)$.

The Fourier coefficients of $f(x)$ are

$$\frac{2h}{\pi\delta}\left(\frac{\sin n\delta/2}{n}\right)^2 e^{-in\xi}.$$

2. The case $1 \leqq p \leqq 2$. One of the defects of Beurling's theorem is that it does not reduce to Wiener's theorem when $p = 2$. Thus while Wiener only requires that S be of measure 0, Beurling demands that S have Hausdorff dimension less than 1.

In terms of Hausdorff α-measures, Beurling's result may be stated as follows.

(2.1) If the α-measure of S is 0 for some $\alpha < 2\frac{p-1}{p}$ then the closure is all of l^p.

One wonders whether this can be replaced by

(2.2) If the α-measure of S is 0 for $\alpha = 2\frac{p-1}{p}$ then the closure is all of l^p.

This is an attractive conjecture and indeed would contain (1.1), (1.2), and (1.3) as special cases. We are, unfortunately, unable to decide whether (2.2) is true or not.

The usual tool, however, for deciding whether the α-measure of a set is 0 is the following implication.

Let S be a closed subset of $[0, 2\pi]$ and let $(a_1, b_1), (a_2, b_2), \cdots (a_n, b_n), \cdots$ be its complementary intervals arranged in non-increasing order of lengths. Call $\Gamma_n = 2\pi - \sum_{k=1}^{n} (b_k - a_k)$.

(2.3) If $\underline{\lim}\, n^{1/\alpha-1} r_n = 0$, then S has α-measure 0.

For $\alpha = 1$, the converse is also true.

Notice, however, that there are in fact denumerable sets S for which r_n goes to 0 arbitrarily slowly so that (2.3) is strictly a one way implication. Nevertheless for the usual 'Cantor-like' sets the condition is both necessary and sufficient and we adopt the following:

Definition. S is said to have strong α-measure 0 if $\lim n^{1/\alpha-1} r_n = 0$.

THEOREM 1. *Let* $1 \leqq p \leqq 2$, $\alpha = 2\dfrac{p-1}{p}$, $f(x) \in l^1$ *and suppose that the set,* S, *of zeros of* $f(x)$, *has strong* α*-measure* 0. *Then the set of multiples,* $P(x)f(x)$, *of* $f(x)$ *by trigonometric polynomials is dense in* l^p.

This theorem, just as Beurling's, does not of course express a nasc. Nevertheless it is best possible, just as is Beurling's, in that it becomes false if $2\dfrac{(p-1)}{p}$ is replaced by any larger number. See Salem [8], whose examples actually have *strong* α-measures 0.

Before we turn to the proof we wish to state a slight generalization.

THEOREM 2. *Let* $1 < p \leqq 2$, $\alpha = 2\,\dfrac{(p-1)}{p}$. *If* S *has strong* α*-measure* 0 *then* S *is a* p*-spanning set.*

LEMMA 3. $\dfrac{\| f \|_p}{\| f \|_2} \leqq A \left[1 + \dfrac{\| f' \|_2}{\| f \|_2}\right]^{(1/p)-(1/2)}$ *for* f *any absolutely continuous function,* A *being an absolute constant.*

Proof. Writing $f(x) = \sum_{-\infty}^{\infty} C_n e^{inx}$, we note that this lemma reduces to a pure inequality for infinite series, namely

$$(1) \qquad \sum_{n=1}^{\infty} | C_n |^p \leqq A \Big(\sum_{n=1}^{\infty} | C_n |^2 \Big)^{\frac{3}{4}p-\frac{1}{2}} \Big(\sum_{n=1}^{\infty} n^2 | C_n |^2 \Big)^{\frac{1}{2}-p/4}$$

Setting $\sum_{n=1}^{\infty} | C_n |^2 = x^2$, $\sum_{n=1}^{\infty} n^2 | C_n |^2 = x^2 y^2$, allows us to rewrite (1) as

$$(2) \qquad \sum_{n=1}^{\infty} | C_n |^p \leqq A x^p y^{1-p/2}.$$

We establish (2) by obtaining the bound $x^p y^{1-p/2}$ for both $\sum_{n\leqq y} | C_n |^p$ and $\sum_{n>y} | C_n |^p$.

$$(3)\qquad \sum_{n\leqq y} | C_n |^p \leqq \Big(\sum_{n\leqq y} | C_n |^2\Big)^{p/2} \cdot \Big(\sum_{n\leqq y} 1\Big)^{1-p/2} \leqq (x^2)^{p/2} y^{1-p/2}.$$

$$(4)\qquad \begin{aligned} \sum_{n>y} | C_n |^p = \sum_{n>y} n^p | C_n |^p 1/n^p &\leqq \Big(\sum_{n>y} n^2 | C_n |^2\Big)^{p/2} \cdot \Big(\sum_{n>y} n^{-2p/2-p}\Big)^{1-p/2} \\ &\leqq (x^2 y^2)^{p/2} \cdot A \cdot (y^{1-2p/2-p})^{1-p/2} = A x^p y^{1-p/2}. \end{aligned}$$

(3) and (4) yield (2) and the proof is complete.

Now choose an n for which $r_n < \epsilon n^{1-1/\alpha}$, define $\rho(x) =$ distance from x to the complement of $\bigcup_{k\leqq n} (a_k, b_k)$ and finally set

$$\Phi(x) = \mathrm{Max}[1 - n^{1/\alpha}\rho(x)/\epsilon, 0]$$

LEMMA 4. $\| \Phi(x) \|_p \leqq A\epsilon^{1-1/p}$.

Proof. Direct application of Lemma 3. We have

$$\| \Phi \|_2 \leqq [\textstyle\int \Phi^2(x)\, dx]^{\frac12} \leqq (r_n + 2n \cdot \epsilon \cdot n^{-1/\alpha})^{\frac12} \leqq \sqrt{3\epsilon}\, n^{\frac12 - 1/2\alpha},$$

Also

$$\| \Phi' \|_2 \leqq (\textstyle\int \Phi'^2(x)\, dx)^{\frac12} \leqq (2n \cdot n^{1/\alpha}/\epsilon)^{\frac12} = \sqrt{2/\epsilon}\, n^{\frac12 + 1/2\alpha}.$$

Hence, by Lemma 3,

$$\| \Phi \|_p \leqq A (\epsilon n^{1-1/\alpha})^{\frac12 - 1/2p} (n^{1+1/\alpha}/\epsilon)^{1/2p - \frac12} = A\epsilon^{1-1/p}.$$

The functions $1 - \Phi(x)$ are Lip 1 and vanish on S, so that an immediate application of Lemma 2 proves Theorem 2.

We now wish to prove a theorem which complements some result of Salem (see [8]). Namely we exploit the connection between closure properties in l^p and existence of functionals in the conjugate space l^q. (Compare Herz [4]).

THEOREM 3. *Let S have strong β-measure 0, and let $d\mu$ be a non-trivial measure whose support lies in S, then* $\sum_n | \int e^{inx} d\mu |^{2/\beta} = \infty$.

Proof. Suppose the series converged so that the Fourier-Stieltjes coefficients of $d\mu$ were $l^{2/\beta}$, then $L(f) = \int f\, d\mu$ would be a non-trivial bounded linear functional on $l^{2/2-\beta}$. But, since the support of $d\mu$ lies in S, L annihilates the ideal of all l^1 functions vanishing in S. Applying Theorem 2 with

$p = 2/2 - \beta$, then, we find that L annihilates a dense subset of $l^{2/2-\beta}$. This contradicts the fact that L is non-trivial and the proof is complete.

We now give another kind of condition on sets, S, which insures that they be p-spanning sets, this condition being related to the notion of a 'base.'

Definition. We say that S is a base of order k if the set $k \times S$, of all numbers of the form $x_1 + x_2 + \cdots x_k$ with each $x_i \in S$, has positive measure.

We say that S is a base if S is a base of order k for some k.

THEOREM 4. *If S is not a base of order k and if $p = 2k/2k - 1$ then S is a p-spanning-set.*

Leaving aside the proof of Theorem 4, for the moment, we wish first to draw some corollaries. In particular, one trivial consequence is

COROLLARY 1. *If S is not a base, then S is a p-spanning-set for all $p > 1$.*

This leads to the following construction: Let K denote the set of all integers in the intervals $[2^k, 2^k(1 + 1/k)]$. Now define S to be the set of all numbers x such that the m-th binary digit of x is 0 for all $m \in K$. S is obviously not a base since no amount of 'carrying' can fill up the blocks of 0's, that is to say if the $x_i \in S$ then $x_1 + \cdots x_n$ still has its m-th binary digit 0 for all m with $2^k \leqq m \leqq 2^k(1 + 1/k) - n$, and so the totality of the numbers $x_1 + \cdots x_n$ has measure 0.

Now consider the set S_ϵ constructed in the same way with the intervals $[2^k, 2^k(1+\epsilon)]$. Clearly S_ϵ can be completely covered by a finite number of translates of the set S (Indeed if the intervals of K corresponding to $k < 1/\epsilon$ are deleted, then the resulting set covers S_ϵ).

It is a simple exercise, however, to show that the dimension of S_ϵ is $\dfrac{1-\epsilon}{1+\epsilon}$ and so we conclude that the dimension of S is 1.

We may sum up these observations in

COROLLARY 2. *There exists a set of dimension 1, which is a p-spanning-set for all $p > 1$.*

Finally we give the

Proof of Theorem 4. Suppose that S is not a $\dfrac{2k}{2k-1}$ spanning-set, then we can find an ideal I with $S_I = S$, and a bounded linear functional, L, on $l^{2k/(2k-1)}$ which annihilates I and is such that $L(1) = 1$.

Reference [6] gives a representation formula for all such functionals L, and this may be put in the following form:

LEMMA 5. *There exists a summable function* $\Phi(x)$ *with* $\Phi'(x)=1$ *for all* $x \notin S$, *such that, for all trigonometric polynomials*, $P(x)$,

$$L(P(x)) = \frac{1}{2\pi}\int_0^{2\pi}[P'(x)\Phi(x)+P(x)]dx.$$

Since, in our present case, L is a bounded linear functional on $l^{2k/(2k-1)}$ it follows that $\{L(e^{inx})\}\in l^{2k}$. Thus if $\{c_n\}$ are the fourier coefficients of $\Phi(x)$, then

$$\sum |nc_n|^{2k} < \infty. \tag{5}$$

If Φ^{*k} denotes the result of convolving Φ with itself k times then Φ^{*k} has fourier coefficients $\{c_n{}^k\}$ and so (5) tells us that the k-th derivative, $(\Phi^{*k})^{(k)}$, exists almost everywhere as an L^2 function with fourier coefficients $(inc_n)^k\}$. It follows, in particular, that

(6) The mean value of $(\Phi^{*k})^{(k)}$ is 0.

The statement (6) will now be contradicted by the following

LEMMA 6. *If* $m \leqq k$, *then* $(\Phi^{*m})^{(m)}=1$ *for all* $x \notin m \times S$. [*In particular* $(\Phi^{*k})^{(k)}=1$ a. e.]

Proof. Induction on m. By Lemma 5 it holds for $m=1$. Suppose that it holds for $m-1$, and that $m \leqq k$. [Also note that all the $m \times S$ are closed sets of measure 0.]
Now

$$\Phi^{*m}(x) = \frac{1}{2\pi}\int_0^{2\pi}\Phi(x-t)\Phi^{*(m-1)}(t)\,dt \tag{7}$$

Let $I_1, I_2, I_3, \cdots$ denote the complementary intervals to $(m-1)\times S$. Since $(m-1)\times S$ has measure 0, (7) may be rewritten as

$$\Phi^{*m}(x) = \frac{1}{2\pi}\sum_{\nu=1}^{\infty}\int_{I_\nu}\Phi(x-t)\Phi^{*(m-1)}(t)\,dt \tag{8}$$

Let x vary over a fixed interval, I, in the complement of $m \times S$. Formal differentiation of (8) leads to

$$(\Phi^{*m}(x))^{(m)} = \frac{1}{2\pi}\sum_{\nu=1}^{\infty}\int_{I_\nu}\Phi'(x-t)\,(\Phi^{*(m-1)}(t))^{(m-1)}dt \tag{9}$$

and since $x-t \notin S$ this in turn is equal, term by term, to

$$\frac{1}{2\pi}\sum_{\nu=1}^{\infty}\int_{I_\nu} 1\cdot 1\,dt.$$

This latter series, of course, converges *uniformly* over the interval, I, and so the formal differentiation is justified. Since the series converges to 1 we conclude that

$$(\Phi^{*m}(x)^{(m)} = 1 \text{ for all } x \in I,$$

and since I was arbitrary, the induction is complete.

This proves Theorem 4.

Before closing this section we wish to point out that Theorem 4 includes Theorem 2 for those p of the special form $\frac{2k}{2k-1}$. This fact rests on the simple observation that a set having strong $1/k$-measure 0 is automatically not a base of order k.

By contrast we mention the famous example, due to Eggleston (see [2]), of a set of Hausdorff dimension 0 which is in fact a base of order 2.

3. The case $\infty > p > 2$. The search for p-spanning-sets in this case seems extremely difficult and we cannot at present even match the questionable precision of the last section. There are, however, two or three fairly trivial observations we can make at the outset.

3.1. A set of 0 measure is a p-spanning-set for any $p > 2$.

3.2. A set containing an interval is not a p-spanning set for any p.

3.3. A set, S, is a p-spanning-set if and only if S supports no non-trivial function in l^q (here, as always, q denotes the conjugate index to p, i. e. $q = \frac{p}{p-1}$).

Clearly $3.3 \Rightarrow 3.1$. Again, since any interval supports a 'triangle' function (which is in fact in l^1), we see that $3.3 \Rightarrow 3.2$. Thus we need only give the (standard duality) argument which proves 3.3.

So let I be a closed ideal in l^1. Once again, by Lemma 1, I contains all Lip 1 functions vanishing on S, and so it contains a single function $G(x)$ whose set of zeros is precisely S.

Now suppose that L is non-trivial bounded linear function on l^p which annihilates I. We know that $(l^p)^* = l^q$ and this tells us that there exists a non-trivial $f(x) \in l^q$ such that $L(g) = \int_0^{2\pi} g(x)f(x)\,dx$ for all $g \in l^1$. Thus $\int_0^{2\pi} g(x)f(x)\,dx = 0$ for all $g \in I$ and in particular $\int_0^{2\pi} e^{inx}G(x)f(x)\,dx = 0$

for all It follows that $G(x)f(x) = 0$, and this is precisely the statement that S supports $f(x)$. Conversely, if S supports a non-trivial function $f(x) \in l$ [illegible]en the functional defined by

$L($ [illegible] $= \int_0^{2\pi} g(x)f(x)\,dx$ for all $g \in l^1$ extends to a non-trivial bounded linear functional on l^p which annihilates I.

These two remarks, of course, complete the proof of 3.3.

We see, then, that our present problem is equivalent to what, in more common Fourier series language, might be termed the uniqueness problem for l^q series. Let us make the following

Definition. We call a closed set S an l^q-multiplicity-set if there exist $\{c_n\}$ not all 0 such that $\sum |c_n|^q < \infty$ and such that $\sum c_n e^{inx}$ converges[8] to 0 for all $x \in S$.

Any closed set which is not an l^q-multiplicity-set we call an l^q-uniqueness-set.

In terms of this definition, 3.3 becomes

3.4. A set is a p-spanning-set if and only if it is an l^q-uniqueness set.

The main results of this section state that there exist p-spanning-sets of positive measure, but that sets which are sufficiently 'well-packed' are not p-spanning-sets.

THEOREM 5. *Let $p > 2$, $\alpha < 2\pi$. There exist p-spanning-sets which have measure $> \alpha$.*

Our next theorem involves the same considerations as were made in Section 2. Namely: If S is a closed set, let $(a_1, b_1), (a_2, b_2), \cdots$ denote its complementary intervals arranged in non-increasing order of length.

This time call $R_n(S) = \sum_{k>n} (b_k - a_k)$.

In general $R_n(S) \to 0$ and the rate at which it does so expresses how 'well packed' the set S is (If, for example, S consists of finitely many intervals, then $R_n(S)$ is eventually equal to 0.).

THEOREM 6. *Let $p > 2$, $\beta > p/2 - 1$, S be a closed set of positive measure. If $R_n(S) = 0_n{}^{-\beta}$ then S is not a p-spanning-set.*

---

[8] If $f(x)$ is a summable function which vanishes on an open set, then its Fourier series converges to 0 on that set. This is simply the observation that convergence is a local property.

The preceding theorem is unimprovable, at least for $p \leqq 4$, in that it becomes false if $p/2-1$ is replaced by any smaller number. [This will be shown at the conclusion of the proof of Theorem 5.]

Theorems 5 and 6 also have 'dual' statements which we can express via 3.4. These are the following:

THEOREM 5′. *Let $1<q<2$, $\alpha<2\pi$. There exist l^q-uniqueness-sets which have measure $>\alpha$.*

THEOREM 6′. *Let $1<q<2$. If S is a closed set of positive measure and if $R_n(S)=0_n{}^{-\beta}$ for some $\beta>\dfrac{2-q}{2q-2}$ then S is an l^q-multiplicity-set.*

We will prove Theorems 5 and 6′, both in a fairly constructive way. Ineed, in order to establish Theorem 6′ we will actually prove

THEOREM 7′. *Let $1<q<2$. If S is a closed set for which $R_n(S)=0_n{}^{-\beta}$, with $\beta>\dfrac{2-q}{2q-2}$, then the 'characteristic function' of S is an l^q function.*

It will prove useful for us to digress a bit with the following number-theoretic considerations, which may also be of independent interest.

Let K be a set of k integers $(\bmod N)$ and denote by $r(n)=r(n;K;N)$ the number of distinct solutions to the congruence

$$n=s+t(\bmod N), s\in K, t\in K.$$

Now if $k=N$, so that K consists of all the integers $(\bmod N)$ the value of $r(n)$ is constantly equal to N, independent of the value of n. On the other hand, as some results of Erdös and Turán (see [3]) show, if k is of the same order as (or of lower order than) $\sqrt{N}$, then, whatever the set K, $r(n)$ must be 'significantly' non-constant. We shall prove, however, that if k is at all of higher order than $\sqrt{N}$, then a K exists which makes $r(n)$ *almost* constant.

The précise statement follows.

LEMMA 7. *There exists a set, K, having k elements, and for which*

$$\frac{1}{N}\sum_{n=1}^{N}(r(n)-c)^2\leqq 5c, \text{ where } c=\frac{k^2}{N}.$$

[In particular if k is large compared to $\sqrt{N}$ then $r(n)$ has a standard deviation which is small compared to its mean.]

Proof. First note that

$$\sum_{n=1}^{N} r(n) = \sum_{n=1}^{N} \sum_{\substack{s,t\in K\\ s+t\equiv n}} 1 = \sum_{s,t\in K} 1 = (\sum_{s\in K} 1)^2 = k^2.$$

so that the required inequality can be written as

$$\mu(K) \leqq k^4/N + 5k^2 \tag{10}$$

where $\mu(K)$ is an abbreviation for $\sum_{n=1}^{N} r^2(n)$.

Next observe that

$$\mu(K) = \sum_{n=1}^{N} (\sum_{\substack{s,t\in K\\ s+t\equiv n}} 1)^2 = \sum_{n=1}^{N} \sum_{\substack{s,t,u,v\in K\\ s+t\equiv u+v\equiv n}} 1 = \sum_{\substack{s,t,u,v\in K\\ s+t\equiv u+v}} 1 \tag{11}$$

If we now form $\sum_{K} \mu(K)$, the sum ranging over all the $\binom{N}{k}$ sets K having precisely k elements, we obtain, from (11),

$$\sum_{K} \mu(K) = \sum_{K} \sum_{\substack{s,t,u,v\in K\\ s+t\equiv u+v}} 1 = \sum_{\substack{s,t,u,v\\ s+t\equiv u+v}} \sum_{\substack{\text{all } K\\ s,t,u,v\in K}} 1 \tag{12}$$

The above inner sum is obviously equal to $\binom{N-i}{k-i}$ where i is the number of *distinct* numbers among s, t, u, v $(i \leqq 4)$. Thus if we denote by λ_i the number of quadruples s, t, u, v for which exactly i of them are distinct and for which $s+t\equiv u+v$ we obtain

$$\sum_{K} \mu(K) = \sum_{i=1}^{4} \lambda_i \binom{N-i}{k-i}. \tag{13}$$

Next a straight forward counting procedure gives

For N even: $\lambda_1 = N$, $\lambda_2 = 2N^2 - N$, $\lambda_3 = 2N^2 - 4N$, $\lambda_4 = N^3 - 4N^2 + 4N$

For N odd: $\lambda_1 = N$, $\lambda_2 = 2N^2 - 2N$, $\lambda_3 = 2N^2 - 2N$, $\lambda_4 = N^3 - 4N^2 + 3N$.

All we need retain, however, are the inequalities

$$\lambda_1 \leqq Nk,\ \lambda_2 \leqq 2N^2,\ \lambda_3 \leqq 2N^3/k,\ \lambda_4 \leqq N^3,$$

together with the obvious inequality

$$\binom{N-i}{k-i} \leqq (k/N)^i \binom{N}{k}$$

and we can conclude that

$$\sum_K \mu(K) \leqq \binom{N}{k}\Big[Nk\frac{k}{N} + 2N^2(k/N)^2 + 2\frac{N^3}{k}(k/N)^3 + N^3(k/N)^4\Big]$$
$$= \binom{N}{k}[5k^2 + k^4/N].$$

It follows, then, that for at least one of these sets K,

$$\mu(K) \leqq k^4/N + 5k^2,$$

And this is exactly the required inequality (10).

We can now translate this lemma into a result on a trigonometric sum.

LEMMA 8. *There exists a set K containing exactly k integers in $[1, N]$ for which*

$$\sum_{n=1}^{N-1} |P(\omega^n)|^4 \leqq 5k^2N, \text{ where } P(t) = \sum_{s\in K} t^s,\ \omega = e^{2\pi i/N}.$$

Proof. Choose the set K given by the previous lemma.

We have

$$\sum_{n=1}^{N} |P(\omega^n)|^4 = \sum_{n=1}^{N} |P^2(\omega^n)|^2 = \sum_{n=1}^{N} \Big|\sum_{m=1}^{N} r(m)\omega^{mn}\Big|^2 = N\sum_{m=1}^{N} r^2(m).$$

But, by Lemma 7, $\sum_{m=1}^{N} r^2(m) \leqq k^4/N + 5k^2$. Thus $\sum_{n=1}^{N} |P(\omega^n)|^4 \leqq k^4 + 5k^2N$ so that $\sum_{n=1}^{N-1} |P(\omega^n)|^4 \leqq 5k^2N$ as required.

In terms of this polynomial, P, we define a function $F(x)$ by

$$F(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

where $c_n = P(\omega^n)/k((\sin \pi n/N)/(\pi n/N))^2$, c_0 being interpreted as 1.

We will now establish the following four properties of $F(x)$.

(14) $F(x)$ is Lip 1

(15) The support of $F(x)$ has measure $4\pi\dfrac{k}{N}$

(16) $\| F(x) - 1 \|_p \leqq 2N^{1/p}k^{-1/2}$ for all p, $2 \leqq p \leqq 4$

(17) $\| F(x) - 1 \|_p \leqq 2N^{1/p}k^{-2/p}$, for $p \geqq 4$.

The function $F(x)$ is easily recognized as the sum of k triangle functions each with base $4\pi/N$ and height N/k. Thus (14) and (15) follow immediately and (16) is established for $p=2$.

By Minkowski's inequality we observe that $\| G \|_p \leqq \| G \|_2^{4/p-1} \| G \|_4^{2-4/p}$ so that (16) need only be established for $p=4$. In this case the inequality states that $\sum_{n=1}^{\infty} | c_n |^4 \leqq 8N/k^2$.

Since $c_n = 0$ for n a multiple of N, we have

$$\begin{aligned}\sum_{n=1}^{\infty} | c_n |^4 &= \sum_{n=1}^{N-1} (| c_n |^4 + | c_{n+N} |^4 + | c_{n+2N} |^4 + \cdots) \\ &\leqq 1/k^4 \sum_{n=1}^{N-1} | P(\omega^n) |^4 (1 + 1/\pi^8 + 1/(2\pi)^8 + 1/(3\pi)^8 + \cdots) \\ &\leqq 8/5k^4 \sum_{n=1}^{N-1} | P(\omega^n) |^4 \leqq 8 \frac{N}{k^2}\end{aligned}$$

(the last inequality following by Lemma 8). This proves (16).

Finally observe that $| c_n | \leqq 1$, so that, for $p \geqq 4$,

$$(\| F(x) - 1 \|_p)^p = \sum_{n \neq 0} | c_n |^p \leqq \sum_{n \neq 0} | c_n |^4 \leqq 16 \frac{N}{k^2}$$

and (17) also follows.

It is clear now, that an appropriate choice of N and k supplies the proof of the following

LEMMA 9. *Let* $p > 2$, $\epsilon > 0$. *There exists a* Lip 1 *function* $F(x)$ *whose support has measure* $< \epsilon$ *and for which* $\| F(x) - 1 \|_p < \epsilon$.

We are now in a position to give the

Proof of Theorem 5. Choose $\epsilon_1, \epsilon_2, \epsilon_3, \cdots$ to be positive numbers for which $\sum \epsilon_i \leqq 2\pi - \alpha$.

For each ϵ_i we choose the corresponding $F(x)$ given by Lemma 9. Calling this function $F_i(x)$ and its set of zeros S_i we choose $S = \bigcap_{i=1}^{\infty} S_i$. Clearly the S_i and so also S are closed sets. Also, since $m(S_i) > 2\pi - \epsilon_i$ it follows that $m(S) > 2\pi - \sum \epsilon_i \geqq \alpha$. But Lemma 9 tells us that $F_i(x)$ is Lip 1, vanishes on S, and $\| F_i(x) - 1 \|_p \to 0$. An application of Lemma 2 completes the proof.

It is now fairly clear why, for $p \leqq 4$, Theorem 6 is best possible. Let us simply be more explicit in our choice of the k, N, S_i etc. above. Indeed let $0 < \gamma < p/2 - 1$ and choose $k = 2^j$, $N = 2^{(\gamma+1)j}$ for $j = j_0, j_0 + 1, \cdots$. By (15) and (16) this choice works. But thus the complementary intervals of S are introduced 2^j at a time having lengths $4\pi 2^{-(\gamma+1)j}$, and it follows that $R_n(S) = 0n^{-\gamma}$.

Finally we present the

Proof of Theorem 7. We prove instead that the characteristic function of T, the complement of S, is l^q. Let T_N be the union of the first N complementary intervals and let $\{a_n(N)\}$, $\{b_n(N)\}$, $\{c_n\}$ denote respectively the fourier coefficients of the characteristic functions of T_N, $T - T_N$, T.

$$c_n = a_n + b_n. \tag{18}$$

Since, by hypothesis, $m(T - T_N) \leqq AN^{-\beta}$, we may add

$$\sum | b_n |^2 \leqq AN^{-\beta}. \tag{19}$$

To estimate the a_n we now prove

LEMMA 10. *If $\Phi(x)$ is the characteristic function of the union of N intervals, $\{a_n\}$ are the fourier coefficients of $\Phi(x)$, and j is any positive integer then*

$$\sum_{2^{j-1} \leqq |n| < 2^j} | a_n |^2 \leqq N 2^{-j}.$$

Proof.

By Parseval's theorem we have

$$4 \sum | a_n |^2 \sin^2 \frac{\pi n}{2^{j+1}} = \frac{1}{2\pi} \int_0^{2\pi} | \Phi(x + \pi/2^j) - \Phi(x) |^2 \, dx. \tag{20}$$

Now $\Phi(x + \pi/2^j) = \Phi(x)$ except possibly in $2N$ intervals of length $\pi/2^j$, so that (20) becomes

$$4 \sum | a_n |^2 \sin^2 \frac{\pi n}{2^{j+1}} \leqq N 2^{-j}. \tag{21}$$

Also, for $2^{j-1} \leqq | n | < 2^j$, $\sin^2 \frac{\pi n}{2^{j+1}} \geqq \frac{1}{2}$ and the result follows. Combining (18), (19) and Lemma 10, we obtain

$$\sum_{2^{j-1} \leqq |n| < 2^j} | c_n |^2 \leqq A (N^{-\beta} + N 2^{-j})$$

or setting $N = [2^{j/1+\beta}]$,

$$\sum_{2^{j-1} \leqq |n| < 2^j} | c_n |^2 \leqq A 2^{-j\beta/1+\beta},$$

and an application of Minkowski's inequality gives

$$\sum_{2^{j-1}\leqq|n|<2^j} |c_n|^q \leqq A2^{j(1-\frac{2\beta+1}{2\beta+2}q)} \tag{22}$$

But $\dfrac{2\beta+1}{2\beta+2}q>1$, since $\beta>\dfrac{2-q}{2q-2}$, and so summing (22) over all j we obtain

$$\sum |c_n|^q<\infty,$$

and the proof is complete.

4. The case $p=\infty$. This situation requires certain modifications before a feasible problem can be realized. l^∞ is not a separable space and so, in particular, the trigonometric polynomials are not dense. Again the closure of translates of a single l^1 function can never be dense, and in fact this closure contains only series whose coefficients $\rightarrow 0$. These ills are all remedied by considering, instead of l^∞, its subspace l_0^∞.

In this case the necessary and sufficient condition that S be a spanning set is simply that S contain no intervals.

We will actually prove somewhat more. Indeed let $\{c_n\}\in l_0^\infty$, we denote by the translates of $\{c_n\}$ all the sequences $\{d_n\}$ where for some integer k, $d_n\equiv c_{n+k}$, and we say that the translates span l_0^∞ if the linear combinations of these $\{d_n\}$ are dense in l_0^∞.

THEOREM 8. *The translates of $\{c_n\}$ span l_0^∞ if and only if there is a dense subset of $(0,2\pi)$ on which the formal series $\sum_{-\infty}^{\infty} c_n e^{inx}$ fails to converge*[4] *to* 0.

Proof. The following two statements are equivalent for $\{c_n\}\in l_0^\infty$

(23) $\sum_{-\infty}^{\infty} c_n e^{inx}$ converges to 0 throughout (a,b)

(24) $\sum_{-\infty}^{\infty} c_n \left(\dfrac{\sin n\epsilon/2}{n\epsilon/2}\right)^2 e^{inx}=0$ throughout $(a+\epsilon,b-\epsilon)$ for every ϵ, $0<\epsilon<\dfrac{b-a}{2}$.

Indeed this is the starting point for Riemann's theory of trigonometric series (see [10]).

---

[4] We mean by the statement '$\sum_{-\infty}^{\infty} a_n$ converges to s' that $\lim_{N\rightarrow\infty} \sum_{-N}^{N} a_n=s$.

Suppose then that our formal series $\sum_{-\infty}^{\infty} c_n e^{inx}$ did converge to 0 in some interval (a, b). Thus the same is true for $\sum_{-\infty}^{\infty} d_n e^{inx}$ where $\{d_n\}$ is any translate of $\{c_n\}$. By (24), then, the bounded linear functional, L, defined by

$$L(\{a_n\}) = \sum a_n \left(\frac{\sin n \dfrac{b-a}{5}}{n \dfrac{b-a}{5}} \right)^2 e^{in\frac{a+b}{2}}$$

annihilates all translates of $\{c_n\}$ so that the translates of $\{c_n\}$ do *not* span l_0^∞.

Conversely if the translates of $\{c_n\}$ do not span l_0^∞ then there exist non-trivial bounded linear functionals which annihilate the translates of $\{c_n\}$. But $(l_0^\infty)^* = l^1$, and the totality of such functionals is thus a closed ideal, I, in l^1. Let (a, b) be an interval which does not intersect S_I. For any ξ, ϵ with $\xi \in (a + \epsilon, b - \epsilon)$, $0 < \epsilon < \dfrac{b-a}{2}$, the function

$$f(x) = \sum \left(\frac{\sin n\epsilon/2}{n\epsilon/2} \right)^2 e^{-in\xi} e^{inx}$$

is a triangle function supported entirely in (a, b). Thus $f(x) \in I$ by Lemma 1 and so the corresponding linear functional L_f annihilates $\{c_n\}$. This means that $\sum c_n \left(\dfrac{\sin n\epsilon/2}{n\epsilon/2} \right)^2 e^{in\xi} = 0$, so that (24) is verified. Thus (23) holds, which means that $\sum c_n e^{inx}$ converges to 0 throughout the interval (a, b).

The proof is complete.

The same proof extends to give

THEOREM 9. *Let* $\mathcal{S}$ *be a closed translation invariant subspace of* l_0^∞. $\mathcal{S}$ *is a proper subspace if and only if there exists an interval* (a, b), $0 < a < b < 2\pi$, *such that* $\sum_{-\infty}^{\infty} c_n e^{inx}$ *converges to* 0 *for all* $x \in (a, b)$ *and all* $\{c_n\} \in \mathcal{S}$.

As a corollary to this theorem we have

COROLLARY. *Let* S *be a closed set containing no intervals and choose any* $\epsilon > 0$. *There exists a* C^∞ *function which is equal to* 1 *throughout* S *and whose fourier coefficients all have moduli less than* ϵ. *(If* S *contained an interval this would no longer be true).*

Proof. Let $f(x) = \sum c_n e^{inx}$ be a C^∞ function which vanishes precisely on S. By Theorem 8, the translates of $\{c_n\}$ span the sequence $\{\delta_{n0}\}$ and this means exactly that, for some trigonometric polynomial, $P(x)$, the function $1 - P(x)f(x)$ is small in the l_0^∞ norm. Since this norm is the maximum of the moduli of the fourier coefficients, the function $1 - P(x)f(x)$ serves to prove the corollary.

An explicit constructive proof of this corollary seems very difficult to give. If it could be given, however, it might shed considerable light on the spanning problem for $p > 2$.

YESHIVA UNIVERSITY.

REFERENCES.

[1] A. Beurling, "On a closure problem," *Arkiv för Matematik*, 1950, pp. 301-303.

[2] H. G. Eggleston, "Note on certain s-dimensional sets," *Fundamenta Mathematicae*, vol. 36 (1949), pp. 40-43.

[3] P. Erdos and P. Turan, "A problem of Sidon in additive number theory," *Journal of the London Mathematical Society*, vol. 16 (1941), pp. 212-215.

[4] C. S. Herz, "A note on the span of translations in L^p," *Proceedings of the American Mathematical Society*, vol. 8 (1957), pp. 724-727.

[5] W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton Press, 1941.

[6] D. J. Newman, "Some results in spectral synthesis," *Duke Mathematical Journal*, vol. 27 (1960), pp. 359-362.

[7] H. Pollard, "The harmonic analysis of bounded functions," *Duke Mathematical Journal*, vol. 20 (1953), pp. 499-512.

[8] R. Salem, "On singular monotonic functions . . .," *Arkiv för Matematik*, 1950, pp. 353-373.

[9] N. Wiener, *The Fourier Integral*, Dover Publications, N. Y.

[10] A. Zygmund, *Trigonometric Series, Vol. I*, Cambridge Press, 1959, Chapter 9.

TWO FUNDAMENTAL THEOREMS ON DEFORMATIONS OF POLARIZED VARIETIES.

By T. MATSUSAKA and D. MUMFORD.[1]

Introduction. In contrast to the theory of moduli of curves, the global theory of moduli of higher dimensional varieties—with the exception of Abelian varieties—is largely unexplored. The work of the authors and of others[2] has begun at least to clarify the problem, and to pose some plausible conjectures. One thing that is clear, however, is that there is a complexity here of a higher order of magnitude from that encountered for curves. The purpose of the present article is to present two results of a qualitative nature that limit the degree of possible complexity of various sought for varieties or scheme of moduli. The first result of ours asserts that two non-singular projective varieties with polarizations, which are isomorphic as polarized varieties, remain isomorphic after specializations over a discrete valuation-ring, whenever they remain non-singular polarized varieties and at least one of them is non-ruled (cf. Th. 2). The second asserts that a set of non-singular polarized surfaces, which are deformations of each other, can be realized as an algebraic family (i. e. a finite union of an irreducible algebraic family) of non-singular projective surfaces in a projective space, if their ranks are bounded; and, in fact, the set of non-singular surfaces with non-degenerate divisors with a given Hilbert polynomial and of any characteristic can be realized as an algebraic family over the ring of integers. From this, it can be shown that the variety of moduli of such surfaces, which are not ruled, is a finite union of Q-varieties, which will be discussed in a near future.

In Chapter I, we shall settle the first result we mentioned. In Chapter II, we give an estimation for $l(X)$ when X is a non-degenerate divisor on a projective variety. Our second main theorem will be settled in Chapter III, as well as in Chapter IV, under slightly different technique. In the first three Chapters, essentially the terminonolgy and conventions of Weil's book [18] are followed. In Chapter IV, because of the nature of the technique which

---

Received October 25, 1963.

[1] This work was done while the first named author was supported by the N. S. F. and the second named author was supported by the Sloan Foundation, and the Army Research Office (Durham).

[2] Cf. [5], [6], [8], [9], [10], [11].

are followed, essentially Grothendieck's terminology and conventions in [2] are followed. However, in order to keep the uniformity, the word *"ample"* (resp. *"non-degenerate") is used for "very ample"* (resp. *"ample"*) in the sense of Grothendieck.

By a *specialization* of a variety or a cycle, we understand a reduction of such over a discrete valuation-ring (cf. [17]). For the theorem of Riemann-Roch in general, we follow quite often the sheaf-theoretic terminology which can be found in [15] and [22]. Let V be a normal variety and M a finitely generated module of functions on V. When $Y=\inf_{g\in M}(\mathrm{div}(g))$, the set $\Lambda(M)$ of V-divisors $\mathrm{div}(g)-Y$, $g\in M$, is called the *reduced linear system* determined by M. When F is any positive V-divisor, $\Lambda(M)+F$ is called a *linear system*. Assume that V is complete. When X is a V-divisor, the set $L(X)$ of functions g on V such that $\mathrm{div}(g)+X\succ 0$ forms a finite dimensional vector space ([18], App. 1, Th. 3). We denote by $\Lambda(X)$ the set of positive V-divisors which are linearly equivalent to X, and call it the *complete linear system* determined by X. We denote by $|X|$ the support of X. We have $\Lambda(X)=\Lambda(L(X))+F$, where $F=X+\inf_{g\in L(X)}(\mathrm{div}(g))$. We denote by $l(X)$ the dimension of $L(X)$. When V is a projective variety, we denote by $\mathfrak{o}_V$ the sheaf of local rings on V, the defining sheaf of functions on a scheme V. If X is a Cartier divisor on V, we denote by $\mathfrak{L}(X)$ the corresponding invertible sheaf. With this sheaf theoretic notations, $H^0(V,\mathfrak{L}(X))=L(X)$ when V is normal. Moreover, when V is a non-singular projective surface, $H^2(V,\mathfrak{L}(X))$ is isomorphic to the dual of $H^0(V,\mathfrak{L}(K(V)-X))$ and $\dim H^1(V,\mathfrak{L}(X))=s(X)$ is the *superabundance* of X. When there is no danger of confusion, we write $H^i(\mathfrak{L}(X))$ for $H^i(V,\mathfrak{L}(X))$.

Chapter I.

THEOREM 1.[3] *Let V be a complete abstract variety, W an abstract variety and T a birational correspondence between V and W. Let $\mathfrak{o}$ be a discrete valuation-ring with the quotient field k, such that V, W and T are defined over k. Let (V',W',T') be a specialization of (V,W,T) over $\mathfrak{o}$ and assume that V', W' are abstract varieties and that V' is complete. When W' is not a ruled variety, there is a component T'' of T' with the coefficient* 1 *in T' such that T'' is a birational correspondence between V' and W' and that $\mathrm{pr}_i(T'-T'')=0$ for $i=1,2$.*

Proof. From the compatibility of specializations with the operation of algebraic projection (cf. [17]), we see that T' has a component T'' with the following properties: (a) $\mathrm{pr}_2 T'' = W'$; (b) the coefficient of T'' in T' is 1; (c) $\mathrm{pr}_2(T' - T'') = 0$. Let $\mathfrak{p}$ be the maximal ideal of $\mathfrak{o}$ and κ the residue field of $\mathfrak{o}$ with respect to $\mathfrak{p}$. Let (x') be a generic point of a representative of W' over κ. Then, there is a representative (x) of a generic point of W over k such that (x') is a specialization of (x) over $\mathfrak{o}$, over

---

[3] This theorem was pionted out to us by M. Artin.

$$(V, W, T) \xrightarrow{\mathfrak{o}} (V', W', T')$$

(cf. [17], Th. 7). Let R_v be the specialization-ring of the specialization $(x) \xrightarrow{\mathfrak{o}} (x')$ in $k(x)$. Then R_v is a discrete valuation ring of $k(x)$ (cf. [17], Prop. 5 and Th. 15). Hence, it determines a valuation v of $k(x)$. Let $Q \times (x)$ be a generic point of T over k. Since V and V' are complete, there is at least one representative (y) of Q such that the coordinates y_i of (y) are in R_v, that (y') is a representative of Q' if $(Q \times (x), (y)) \xrightarrow{\mathfrak{o}} (Q' \times (x'), y'))$ and that $Q' \times (x')$ is contained in $|T'|$. When that is so, $Q' \times (x')$ is contained in T''; in fact, it is a generic point of T'' since $\mathrm{pr}_2 \colon T'' \to W'$ is birational and (x') is a generic point of W' over $\bar{\kappa}$. It follows that Q' is a generic point of the projection A of T'' on V' over $\bar{\kappa}$. Let R be the specialization ring of $(y) \xrightarrow{\mathfrak{o}} (y')$. Then, the valuation v is a prime divisor of R in the sense of Abhyankar, and W' is a ruled variety over A unless $A = V'$ (cf. [1], Prop. 3). Therefore, $A = V'$. When that is so, T'' is a birational correspondence between V' and W', which can be seen easily, using the compatibility of specializations with the operation of intersection-product (cf. [17]).

THEOREM 2. *Let $\mathfrak{o}$ be a discrete valuation-ring with the quotient field k; let V and W be non-singular projective varieties, defined over k, and T the graph of an isomorphism, defined over k, between V and W. Let X (resp. Y) be a non-degenerate divisor on V (resp. W), both rational over k, such that $Y = T(X)$. Let $(V, W, X, Y, T) \xrightarrow{\mathfrak{o}} (V', W', X', Y', T')$ and assume that V', W' are non-singular and that X' (resp. Y') is also non-degenerate on V' (resp. W'). Then T' is the graph of an isomorphism between V' and W', if one of the V', W' is not ruled.*

Proof. By Theorem 1, we have $T' = T'' + T^*$, where T'' is a birational correspondence between V' and W', and $\mathrm{pr}_l T^* = 0$ for $l = 1, 2$. Let $F_1, \cdots, F_t$ be the projections of the components of T^* on V'. Note that none of the F_i is 0-dimensional: for if F_i were 0-dimensional, the corresponding component of the n-dimensional cycle T^* would have to be of the form $F_i \times W'$, and this contradicts $\mathrm{pr}_2 T^* = 0$. If X'_m is a divisor in $\Lambda(mX')$, then T' and $X'_m \times W'$ intersect properly if and only if $|X'_m| \not\supset F_i$ for any i. Let U be the set of such divisors X'_m. For every such X'_m, $T'(X'_m)$ is defined. The Chow-variety of U, i.e. the set of Chow-points of members of U, is an open subset of that of $\Lambda(T'(X'_m))$. When U is not empty, the mapping $X'_m \rightarrow T'(X'_m)$ defines, as is well-known, an injection of U into $\Lambda(T'(X'_m))$; and, as a matter of fact, defines an injective linear rational map of the Chow variety of U into that of $\Lambda(T'(X'_m))$ cf. [18], Chap. IX, Th. 3 and [18], Chap. VIII, Th. 4). Now assume that (a) $T'(X'_m) \sim mY'$ and (b) $l(mX') = l(mY')$ for large m.

Suppose that P' is a point of V' and let $\Lambda(mX')_{P'}$ be the linear subsystem of divisors which pass through P'. For sufficiently large m, mX' is ample, hence P' is the only base point of $\Lambda(mX')_{P'}$, hence $\Lambda(mX')_{P'} \cap U$ is not empty. Then the set A of divisors $T'(Z')$, $Z' \in \Lambda(mX')_{P'} \cap U$, consists of divisors passing through every point Q' such that $P' \times Q' \in |T'|$. If there were more than one such Q', the closure of the Chow-variety of A is at least of co-dimension 2 in that of $\Lambda(mY')$, since mY' is ample for sufficiently large m. On the other hand, its co-dimension has to be 1 as the closure of the image of the Chow-variety of $\Lambda(mX')_{P'} \cap U$ by the injective rational map, since $l(mX') = l(mY')$. Hence $T^* = 0$ and T' is single-valued on the points of V', hence everywhere regular by Zariski's Main Theorem. Similarly T'^{-1} is everywhere regular.

To prove (a) and (b), note that $p_a(mX) = p_a(mY)$ for all integers m. Hence $p_a(V') = p_a(V) = p_a(W) = p_a(W')$; and $p_a(mX') = p_a(mX) = p_a(mY) = p_a(mY')$ for all integers m (cf. [13]). It follows that $l(mX) = l(mY) = l(mX') = l(mY')$ for large positive integer m by the theorem of Riemann-Roch (cf. [21]). Thus (b) is satisfied. Now let C and D be the supports of the Chow-varieties of $\Lambda(mX)$, $\Lambda(mY)$. Since the linear equivalence is preserved under specializations (cf. [17]), C', D' will be the supports of the Chow-varieties of $\Lambda(mX')$, $\Lambda(mY')$ for large m, if $(C, D) \xrightarrow{\mathfrak{o}} (C', D')$. Then (a) follows from the compatibility of specializations with the operation of intersection-product and from the invariance of linear equivalence under specializations.

Let V and V' be two complete non-singular polarized varieties (cf. [20]), k a field of definition of V and $\mathfrak{o}$ a discrete valuation ring with the quotient field k. Let X be a polar divisor of W and W, W' the underlying varieties of V, V'. If $(W, X) \xrightarrow{\mathfrak{o}} (W', X')$ and X' is a polar divisor of V', we shall say that V' is a specialization of V over $\mathfrak{o}$. With this definition, de have the following corollary.

COROLLARY 1. *Let $\mathbf{V}$ and $\mathbf{W}$ be two varieties over a discrete valuation ring $\mathfrak{o}$ (i.e. $\mathfrak{p}$-variety in the sense of Shimura; a scheme in the sense of Grothendieck). Let generic fibres V, W of $\mathbf{V}$, $\mathbf{W}$ be non-singular projective varieties, defined over the quotient field k of $\mathfrak{o}$. Let the special fibres V', W' be non-singular projective varieties. Assume that V, W, V', W' are underlying varieties of polarized varieties $\bar{V}$, $\bar{W}$, $\bar{V}'$, $\bar{W}'$ and that $(V, W) \xrightarrow{\mathfrak{o}} (V', W')$ can be extended to $(\bar{V}, \bar{W}) \xrightarrow{\mathfrak{o}} (\bar{V}', \bar{W}')$. Then, when there is an isomorphism $\bar{f}$ between $\bar{V}$ and $\bar{W}$ over k, $\bar{f}$ can be extended to an isomorphism $\mathbf{f}$ of $\mathbf{V}$ and $\mathbf{W}$, if W' is not ruled. Moreover, the graph of $\bar{\mathbf{f}}$ specializes to the graph of an isomorphism f' between $\bar{V}'$ and $\bar{W}'$ over $\mathfrak{o}$.*

COROLLARY 2. *Let V be a projective, non-ruled, non-singular variety with a structure of polarization and G the connected component, containing the identity, of the group of automorphisms of V. Then G is an Abelian variety.*

Proof. The group of automorphisms of V is an algebraic group (cf. [8]). If G is not complete, the graph of an automorphism, corresponding to a suitable element of G, can be specialized, over some field of definition of V, to a $V \times V$-cycle which is not the graph of an automorphism. This is impossible by Theorem 2. Hence G is complete and is an Abelian variety by the theorem of Chevalley (cf. [19], Th. 5).

Chapter II.

Let V^n be a normal projective variety and X a non-degenerate divisor on V. $L(mX)$ defines a projective embedding f_m of V for large m by the definition. Let W^r be a simple subvariety of V and k a common field of definition for W and V, over which X is rational. Then $L(mX)$ has a basis over k (cf. [18], Ch. IX, Cor. 1 of Th. 8). Let $A_1, \cdots, A_r$ be independent generic divisors of $\Lambda(mX)$ over k. Then every component of $W \cap A_1 \cap \cdots \cap A_r$ is simple on V and on W (cf. [18], Ch. V, Th. 1). We set $[W \cdot X^{(r)}] = (1/m^r)\deg(W \cdot A_1 \cdots A_r)$ and $X^{(n)} = [V \cdot X^{(n)}]$. Then $[W \cdot X^{(r)}]$ does not depend upon the choice of independent generic divisors $A_1, \cdots, A_r$. Moreover, it does not depend upon the choice of m, as long as it is sufficiently large, and is a positive integer (cf. Bezout's theorem).

THEOREM 3. *Let V^n be a normal projective variety and X a non-degenerate divisor on V. Then $l(X) \leqq X^{(n)} + n$.*

Proof. If $l(X) \leqq 1$, there is nothing to prove. If $n = 1$, our theorem is an immediate consequence of the theorem of Riemann-Roch. Therefore, we assume that $l(X) > 1$ and that $n > 1$. Let X_0 be a generic divisor of the complete linear system $\Lambda(X)$. Then $X_0 = \sum_1^d Y_i + F$, where F is the fixed component of $\Lambda(X)$ and $Y_i \neq Y_j$ for $i \neq j$, since $\Lambda(X)$ is complete (cf. [18], Ch. IX, Cor. of Th. 15). If $d > 1$, the Y_i are generic divisors of one and the same pencil on V by the theorem of Bertini (cf. [18], Ch. IX, Th. 17). Hence $\dim \Lambda(X) = l(X) - 1 \leqq d$. On the other hand, we get $d \leqq X^{(n)}$ by computing $X^{(n)} = [(\sum_1^d Y_i + F) \cdot X^{(n-1)}]$. Therefore, our theorem is true in this case also.

Assume now that $d = 1$. Then $X_0 = Z + F$, where Z is an absolutely irreducible subvariety of V. Let K be an algebraically closed field, containing k, over which Z and F are rational, and (Z^*, α) a normalization of Z over K. Let m_0 be a positive integer such that mX is ample for $m \geqq m_0$ and X_m a generic divisor of $\Lambda(mX)$ over K for such m. We note that every component of $X_m \cap Z$ is simple both on V and Z and is proper on V. We contend that: (a) When $g \in L(X_{m+1} - X_m)$, $g \to g^* = g \circ \alpha^{-1}$ is a homomorphism of $L(X_{m+1} - X_m)$ into $L(X^*)$, where $X^* = \alpha(Z \cdot (X_{m+1} - X_m))$; (b) The kernel of the above homomorphism is a vector space of dimension 1; (c) X^* is non-degenerate on Z^* and $X^{*(n-1)} \leqq X^{(n)}$. Our theorem will be an immediate consequence of (a), (b), (c). For, we have $l(X^*) \leqq X^{*(n-1)} + (n-1)$ by the induction hypothesis, hence $l(X^*) \leqq X^{(n)} + (n-1)$ by (c), and $l(X) = l(X_{m+1} - X_m) \leqq l(X^*) + 1$ by (a) and (b).

To prove (a), we may assume that $g^* \neq 0$. We first remark the following two facts: (i) If U is a subvariety of Z of co-dimension 1, which is simple both on V and Z, and g' is the function induced on Z by g, then the coefficient of U in $\operatorname{div}(g) \cdot Z$, that of U in $\operatorname{div}(g')$ and that of $\alpha(U)$ in $\operatorname{div}(g^*)$ all coincide; (ii) If W^* is a component of $\operatorname{div}(g^*)_\infty = g^{*-1}(\infty)$, its geometric image W by α^{-1} is a component of $X_{m+1} \cap Z$, and is simple both on V and Z. In fact, U has the same coefficient a in $\operatorname{div}(g) \cdot Z$ as in $\operatorname{div}(g')$ (cf. [18]-IX, Th. 3). Since g^* can be written as $g' \circ \alpha^{-1}$, and since α is biregular along U, it follows that the coefficient of $\alpha(U)$ in $\operatorname{div}(g^*)$ is also a. As for (ii), g^* is not finite along W^* (i. e. at a generic point of W^* over a field of definition of W^*, containing K), and hence, g is not also finite along W. Consequently, $W \subset |g^{-1}(\infty)| = |X_{m+1}|$ and W is a component of $Z \cap X_{m+1}$.

Now let U be a component of $X_m \cap Z$ or of $X_{m+1} \cap Z$. Since

$$\operatorname{div}(g) \cdot Z + Z \cdot (X_{m+1} - X_m) \succ 0,$$

it follows that the coefficient of $\alpha(U)$ in $\operatorname{div}(g^*) + X^*$ is non-negative by (i). Therefore, if $\operatorname{div}(g^*) + X^*$ has a component W^* of negative coefficient, it is a component of $\operatorname{div}(g^*)_\infty = g^{*-1}(\infty)$, which is impossible by (ii).

To prove (b), let g be a function in $L(X_{m+1} - X_m)$ such that $g^* = 0$. Then $\operatorname{div}(g) = W - (X_{m+1} - X_m)$, where W is a positive V-divisor such that Z is a component of it. Since $W \sim X_{m+1} - X_m \sim X$, it follows that $W = Z + F$ and that g is uniquely determined up to a constant factor. (b) is thereby proved.

To prove (c), choose a positive integer $r \geqq m_0$ and identify $\Lambda(rX)$ with the linear system of hyperplane sections of V by means of the embedding f_r. Let s be another large positive integer. Then $sX_r \sim srX_{m+1} - srX_m$ and sX_r, srX_m, srX_{m+1} are sections of V by hypersurfaces of degrees s, sm, $s(m+1)$ respectively, since the linear system of hypersurface sections of a normal projective variety is complete when the degree of hypersurfaces is large enough (Zariski's normalization theorem). Consequently, $sX_r \cdot Z$, $srX_m \cdot Z$, $srX_{m+1} \cdot Z$ are also sections of Z by hypersurfaces of degrees s, sm, $s(m+1)$. When s is chosen large enough so that α is determined by homogeneous functions of homogenity s, $\alpha(sX_r \cdot Z)$, $\alpha(srX_m \cdot Z)$, $\alpha(srX_{m+1})$ are hypersurface sections of Z^* by hypersurfaces of degrees 1, m, $m+1$ respectively. Hence

$$\alpha(sX_r \cdot Z) \sim sr\alpha(Z \cdot (X_{m+1} - X_m)).$$

Thus, $X^* = \alpha(Z \cdot (X_{m+1} - X_m))$ is non-degenerate. $\Lambda(rX)$ is the linear system of hyperplane sections of V. Hence

$$X^{(n)} = (1/r^{n-1})\deg(X) = (1/r^{n-1})\deg(Z+F) \geqq (1/r^{n-1})\deg(Z).$$

$\Lambda(srX^*)$ is the linear system of hyperplane sections of Z^*. Hence $X^{*(n-1)} = (1/(sr)^{n-1})\deg(Z^*)$. But $\deg(Z^*) = s^{n-1}\deg(Z)$ as is well-known and easy to see. (c) is thus proved.

Remark 1. Let $\mathfrak{L}$ be a non-degenerate invertible sheaf (ample invertible sheaf in the sense of Grothendieck) on a projctive variety V. Let d be the leading coefficient of $\chi(\mathfrak{L}^m)$. Then it is easy to deduce that $\dim H^0(\mathfrak{L}) \leqq d + \dim V$ from our theorem. In fact, when (V^*, β) is a normalization of V and X a Cartier divisor on V^* determined by $\mathfrak{L}$, then X is non-degenerate and $d = X^{(n)}$ if $\dim V^* = n$.

Remark 2. In our theorem, we assumed that X is non-degenerate. Assume now that V is a non-singular projective surface and X a V-divisor such that $X^{(2)} > 0$ and $[Y \cdot X] > 0$ for all positive V-divisors Y. Then, we can prove directly that $l(X) \leqq X^{(3)} + 2$. In fact, we may assume, as in the proof of our theorem, that a generic divisor X_0 of $\Lambda(X)$ is of the form $Z + F$, where Z is an irreducible curve. Let Y be a V-divisor such that $Y \sim X$ and that $|Y|$ contains neither Z nor any singular point of Z. Then $[X \cdot Z] = [Y \cdot Z] \leqq X^{(2)}$ and $L(X)$ and $L(Y)$ are isomorphic. As in the proof of our theorem, $L(Y)$ induces on Z a module M' of functions on Z; $\alpha^{-1}(M')$ is then a submodule of $M^* = L(\alpha^{-1}(Z \cdot Y))$ and the kernel of the homomorphism $L(Y) \to M^*$ is a vector space of dimension 1. Hence, we have our inequality by the theorem of Riemann-Roch. Our divisor X is in fact a non-degenerate divisor on V according to [12], and our theorem is available according to this. But using this remark and our Theorem 4, we recover this result.

Chapter III.

Let V be a non-singular projective surface and X, Y two divisors. There is a V-divisor X' such that $X' \sim X$ and that X' and Y intersect properly on V. We denote by $X_\wedge Y$ the intersection-product $X' \cdot Y$, and by $[X \cdot Y]$ the degree of $X_\wedge Y$. When $X = Y$, $[X \cdot X]$ is denoted by $X^{(2)}$. We denote by $K(V)$ a canonical divisor on V and set $p_a(X) = (1/2)[X \cdot (X + K(V))] + 1$. When X is irreducible, $(X + K(V))_\wedge X$ is a canonical divisor $K(X)$ of X and $\deg(K(X)) = 2p_a(X) - 2$ (cf. [16]). When $X = \sum_i a_i X_i$, we have

$$p_a(X) = \sum_i a_i p_a(X_i) + \sum_i (1/2) a_i (a_i - 1) X_i^{(2)} \tag{1}$$
$$+ \sum_{i,j;\, i \neq j} (1/2) a_i a_j [X_i, X_j] - \sum_i a_i - 1.$$

According to the theorem of Riemann-Roch on V, we have

$$l(X) - s(X) + l(K(V) - X) = X^{(2)} - p_a(X) + p_a(V) + 2.$$

1. Denote by Σ the set of pairs (V, X) of a projective non-singular surface V and a V-divisor X satisfying the following conditions.

(I) $[X \cdot Y] > 0$ whenever Y is a positive V-divisor;
(II) $0 < X^{(2)} < c_1$;
(III) $|p_a(X)| < c_2$;
(IV) $|p_a(V)| < c_3$.

In order to simplify further discussions, we assume that the constants c_i $(i > 3)$ which will be introduced are *positive integers, satisfying* $c_i > c_{i-1}$ and depending *only upon* c_1, c_2, c_3.

LEMMA 1. *There are constants* c_4, c_5 *such that* $|[X\cdot K(V)]|<c_4$ *and that* $p_a(mX)>0$, $l(K(V)-mX)=0$, $m^2X^{(2)}-p_a(mX)+p_a(V)+1>0$ *whenever* $(V,X)\in\Sigma$ *and* $m>c_5$.

This is an easy consequence of (I), (II), (III), (IV), the theorem of Riemann-Roch and of the formula (1).

LEMMA 2. *Let* (V,X) *be a member of* Σ *and* $T=\sum_1^t a_iY_i$ *the reduced expression for a member* T *of* $\Lambda(2c_5X)$. *Then, there are constants* c_6, c_7, c_8 *and* c_9 *with the following properties:*

(i) $\sum_1^t a_i<c_6$;

(ii) $|[Y_i\cdot Y_j]|<c_7$;

(iii) $|[K(V)\cdot Y_i]|<c_8; 0\leqq p_a(Y_i)<c_8$;

(iv) The multiplicity of any point on Y_i is at most c_9.

Proof. (i) is a consequence of $\sum_1^t a_i\leqq\sum_1^t a_i[Y_i\cdot X]=[T\cdot X]\leqq 2c_1c_5$. (ii) and (iii) follow from the three inequalities:

(A) $a_iY_i^{(2)}+\sum_{i\neq j}a_j[Y_i\cdot Y_j]=[Y_i\cdot 2c_5X]\leqq\sum a_i[Y_i\cdot 2c_5X]\leqq 4c_5^2c_1$.

(B) $-2\leqq 2p_a(Y_i)-2=[Y_i\cdot(Y_i+K(V))]$.

(C) $\sum a_i[K(V)\cdot Y_i]\leqq 2c_4c_5$.

In fact, (A) gives an upper bound for every $Y_i^{(2)}$. Hence, (B) gives a lower bound for every $[K(V)\cdot Y_i]$. Then (C) gives an upper bound for every $[K(V)\cdot Y_i]$ and (iii) is proved. Returning to (B), we obtain a lower bound for every $Y_i^{(2)}$, and using this, (A) gives upper bounds for all $[Y_i\cdot Y_j]$. This gives (ii), since $[Y_i\cdot Y_j]\geqq 0$ if $i\neq j$. Finally, the arithmetic genus of Y_i is bounded by (ii) and (iii). If the r_{ij} are the multiplicities of the singular points x_{ij} of Y_i, an inequality of Noether (cf. [4]) states

$$\sum_j r_{ij}(r_{ij}+1)/2+p_a(Y_i^*)\leqq p_a(Y_i),$$

where Y_i^* is a non-singular model of Y_i. This gives (iv).

The following lemma is an easy consequence of the generalized Riemann-Roch theorem for curves.

LEMMA 3. *Let* W *be a non-singular surface in a projective space and* Y *a divisor on* W. *Let* C *be an irreducible curve on* W *such that* $[Y\cdot C]>2p_a(C)-2$. *Then* $H^1(\mathfrak{L}(Y)/\mathfrak{L}(Y-C))=0$.

COROLLARY. *Using the same assumptions and notations of our lemma,* $s(Y-C)=s(Y)$ *if and only if*

$$0\to H^0(\mathfrak{L}(Y-C))\to H^0(\mathfrak{L}(Y))\to H^0(\mathfrak{L}(Y)/\mathfrak{L}(Y-C))\to 0.$$

Proof. This is an immediate consequence of our lemma and of an exact sequence $0\to\mathfrak{L}(Y-C)\to\mathfrak{L}(Y)\to\mathfrak{L}(Y)/\mathfrak{L}(Y-C)\to 0$.

LEMMA 4. *Let* $T=\sum_i a_iY_i$ *be a positive divisor on a non-singular projective variety* W, $\Lambda(A)$ *a complete linear system on* W *and assume that* $[(A-T')\cdot Y_i]>2p_a(Y_i)-2$ *for all* i *and for all* T' *such that* $0\prec T'\prec T$. *Then we have* $H^1(\mathfrak{L}(A)/\mathfrak{L}(A-T))=0$.

Proof. If $\sum_i a_i=1$, our lemma follows from Lemma 3. Assume that our lemma has been proved for those positive W-divisors $T''=\sum_i a''_iY_i$ with $\sum_i a''_i<\sum_i a_i$. Set $T'=\sum_i a'_iY_i$ with $a_1-1=a'_1$, $a_i=a'_i$ for $i\geqq 2$. In the exact cohomology sequence of an exact sequence

$$0\to\mathfrak{L}(A-T')/\mathfrak{L}(A-T)\to\mathfrak{L}(A)/\mathfrak{L}(A-T)\to\mathfrak{L}(A)/\mathfrak{L}(A-T')\to 0,$$

we have $H^1(\mathfrak{L}(A)/\mathfrak{L}(A-T'))=0$ by the induction assumption, and $H^1(\mathfrak{L}(A-T')/\mathfrak{L}(A-T))=0$ by our assumption and Lemma 3. Hence we get $H^1(\mathfrak{L}(A)/\mathfrak{L}(A-T))=0$.

COROLLARY 1. *Let* (V,X) *be a member of* Σ *and* $T=\sum_i a_iY_i$ *the reduced expression for a member of* $\Lambda(2c_5X)$. *Set* $T'=\sum_i a'_iY_i$, $U=\sum_i a''_iY_i$ *with* $0\leqq a'_i$, $a''_i\leqq a_i$. *Then, there is a constant* c_{10} *such that*

$$H^1(\mathfrak{L}(2mc_5X-U)/\mathfrak{L}(2mc_5X-U-T'))=0$$

for $m\geqq c_{10}$.

Proof. This follows at once from our lemma, (I) and from Lemma 2.

COROLLARY 2. *There is a constant* c_{11} *such that*

$$c_{11}\geqq s(2mc_5X)\geqq s(2(m+1)c_5X-B)\geqq s(2(m+1)c_5X)$$

whenever $(V,X)\in\Sigma$, $l(2c_5X-B)\geqq 1$ *and* $m\geqq c_{10}$.

Proof. This is an easy consequence of Theorem 3, Corollary 1 above and of Lemma 2. (cf. Remark 2.)

COROLLARY 3. *There is a constant* c_{12} *such that* $\Lambda(2mc_5X)$ *is irreducible (i. e. contains an irreducible curve) whenever* $(V,X)\in\Sigma$ *and* $m>c_{12}$.

Proof. Let T be a member of $\Lambda(2c_3X)$. If Y is a fixed component of $\Lambda(mT)$, we have $l(mT) = l(mT - Y)$. Then $m[T \cdot Y] = s(mT + Y) - s(mT) + p_a(Y) - 1$ by the theorem of Riemann-Roch, which leads to a contradiction if $m > \max(c_{10}, 2c_{11} + c_8 - 1) = c_{11}'$ by the above Corollary 1, (I) and by (iii) of Lemma 2. If $\Lambda(mT)$ is composed of a pencil for $m > c_{11}'$, a generic divisor of $\Lambda(mT)$ can be written as $\sum_1^t T_i$, where the T_i belong to one and the same pencil by the theorem of Bertini. Clearly, we have $\dim \Lambda(mT) \leqq t$ and $[T_i \cdot T_j] \geqq 1$ by (I). Hence $(mT)^{(2)} \geqq t^2$ and $2mc_5 \cdot c_1^{\frac{1}{2}} \geqq t \geqq \dim \Lambda(mT)$. On the other hand,

$$\dim \Lambda(mT) \leqq (2m^2c_5^2c_1 - mc_5c_4) - c_8$$

by (I), Lemma 1 and by the theorem of Riemann-Roch. Our corollary now follows from this easily.

2. When Λ and Λ' are two linear systems on a complete normal variety, the smallest linear system Λ'' containing the divisors $X + X'$, $X \in \Lambda$, $X' \in \Lambda'$, is called the *minimum sum* of Λ and Λ'. Then the following lemma is easy to prove.

LEMMA 5. *Let $\Lambda(C)$ be a non-empty complete linear system on a complete normal variety W. Assume that $\Lambda(C)$ has no base point and that the minimum sum of $\Lambda(C)$ and $\Lambda(mC)$ is complete. Let h_i be a non-degenerate map of W into a projective space determined by $\Lambda(iC)$ for $i = m$, $m+1$. Then there is an isomorphism α between images W_m, W_{m+1} of W by h_m, h_{m+1} such that $h_{m+1} = \alpha \circ h_m$.*

In the following three lemmas, denote by C an irreducible curve on a non-singular projective surface V and $\mathfrak{R}_C$ the intersection of local rings of C at the singular points of C. Using only those functions of C which are in $\mathfrak{R}_C$, we can define linear systems as in the case of normal varieties. Throughout this chapter, *linear systems on curves lying on V are understood in this sense.* By the *degree* of a linear system on C, we understand the degree of a generic divisor of the linear system. The Riemann-Roch theorem on C then states $l(\mathfrak{m}) = \deg(\mathfrak{m}) - p_a(C) + 1 + l(K(C) - \mathfrak{m})$ for a C-divisor $\mathfrak{m}$ (cf. [13], [16]). The following lemma is known as a lemma of Castelnuovo when C is non-singular, which can be proved in the same way as in the ordinary case.

LEMMA 6. *Let Λ' be a linear system on C without base point and Λ a complete non-special linear system on C. Let $\mathfrak{n}'$ be a generic divisor of Λ' and assume that $\Lambda - \mathfrak{n}'$ is non-special and is of degree equal to* $\deg(\Lambda) - \deg(\mathfrak{n}')$. *Then the minimum sum of Λ and Λ' is complete.*

Actually, it is enough to know a special case of this lemma, under an additional assumption that $\deg(\Lambda) - \deg(\mathfrak{n}') \geqq 2p_a(C)$, which makes a proof very easy.

Let Λ be a linear system $\Lambda(M) + F$, where M is a finitely generated module of functions on V. Let C^* be the largest non-singular open subset of C and denote by $*$ the restriction of a C-chain (i.e. a zero-cycle on V whose support is contained in C) to C^*. Assume that C and F intersect properly on V and that every g in M induces a function g' in $\mathfrak{R}_C$ on C. Denote by M' the set of such functions g' and by Λ' the set of C-divisors $(X \cdot C)^* + (C \cdot F)^*$, by taking for X all divisors from $\Lambda(M)$ such that X and C intersect properly on V. Then Λ' is a linear system on C, whose reduced part is determined by M'. We denote Λ' by $\mathrm{Tr}_C\,\Lambda$ and call it the linear system on C induced by Λ. When Λ is a complete linear system $\Lambda(X)$ whose fixed component F satisfies our requirement, it always induces a linear system on C, since there is a V-divisor Z such that $X \sim Z$ and that the support of Z does not contain C and the singular points of C.

LEMMA 7. *Assume that* $C^{(2)} > 0$ *and that* $\Lambda(C)$ *has no base point. If* $s(mC)$ *is a constant for all positive integers* m, *the minimum sum of* $\Lambda(C)$ *and* $\Lambda(mC)$ *is complete for* $m > [C \cdot K(V)] + 4$.

Proof. Let C' be a V-divisor such that $C' \sim C$ and that $|C'|$ contains neither C nor the singular points of C. We have $[C \cdot (mC)] - 2p_a(C) > 0$ when m satisfies our condition. Hence $\mathrm{Tr}_C\,\Lambda(mC')$ is complete by Corollary of Lemma 3. By our assumption, $\mathrm{Tr}_C\,\Lambda(C')$ has no base point. Hence the minimum sum of $\mathrm{Tr}_C\,\Lambda(mC')$ and $\mathrm{Tr}_C\,\Lambda(C')$ is complete for such m by Lemma 6. Thus, the minimum sum of $\Lambda(mC')$ and $\Lambda(C')$ induces on C a complete linear system. Let M be the module generated by $f \cdot g$ with $f \in L(mC')$, $g \in L(C')$. Then M induces on C the module $L((m+1)C' \cdot C)$. When h is a function in $L((m+1)C')$, inducing 0 on C, we have $\mathrm{div}(h) = C + H - (m+1)C'$ with $H \succ 0$. Hence h is in M (cf. [18], Chap. IX, Cor. 2 of Th. 8). Our lemma follows from this at once.

It is not true in general that a complete linear system $\Lambda(\mathfrak{m})$ on C contains a divisor of the same degree as $\mathfrak{m}$, unless $\mathfrak{m}$ is positive. But we have the following.

LEMMA 8. *When* $\deg(\mathfrak{m}) \geqq 2p_a(C)$, $\Lambda(\mathfrak{m})$ *contains a divisor of the same degree as* $\mathfrak{m}$.

Proof. Let $\mathfrak{a}$ be a positive C-divisor such that $\deg(\mathfrak{m}) = \deg(\mathfrak{a}) = m$. Set $p_a(C) = l$ and let k be a field of definition for C over which $\mathfrak{m}$ and $\mathfrak{a}$

are rational. Then, there are two generic divisors $\mathfrak{p}$ and $\mathfrak{q}$ of degrees t over k and a unit f in $\mathfrak{R}_C$ such that $\operatorname{div}(f) = (\mathfrak{m} - \mathfrak{a}) + (\mathfrak{p} - \mathfrak{q})$ (cf. [14], Lemma 3). Since $\mathfrak{a}$ is positive, a generic divisor of $\Lambda(\mathfrak{a})$ over k has the degree m; moreover, we have $l(\mathfrak{a}) \geqq t + 1$. Therefore, there is a unit g in $\mathfrak{R}_C$ and a generic divisor $\mathfrak{a}' = \mathfrak{p} + \mathfrak{b}$ of $\Lambda(\mathfrak{a})$ of degree m such that $\operatorname{div}(g) = \mathfrak{a} - \mathfrak{a}'$. Then $\operatorname{div}(f \cdot g) = \mathfrak{m} - \mathfrak{q} - \mathfrak{b}$ and our lemma is proved.

3. THEOREM 4. *There is a constant* c_{13} *such that* mX *is ample for all* $(V, X) \in \Sigma$, *whenever* $m > c_{13}$.

Proof. In order to prove our theorem, it is enough to prove that m_0X is ample for all (V, X) in Σ, where m_0 is a constant, depending only on c_1, c_2, c_3. In fact, the sum of two ample divisors is also ample; moreover, if Y is an ample divisor on a non-singular projective variety W and $B = B' - B''$, $B' > 0$, $B'' > 0$, is a W-divisor, then, whenever

$$d \geqq \deg(B') \cdot (\deg(W) - 2) + \deg(V) + \deg(B''),$$

$dY + B$ is ample (cf. [18], Chap. IX, Cor., Th. 13). Our theorem follows from these two facts and from Lemma 1 as an easy exercise.

Let T be a member of $\Lambda(2c_5X)$ and Z an irreducible member of $\Lambda(c_{12}T)$ (cf. Cor. 3 of Lemma 4). Then we have

(a) $[rZ \cdot (2tZ + iT)] > 2p_a(rZ)$ *for* $0 < r \leqq t$ *and for all* $i \geqq 0$. In fact, this follows from (II), Lemma 1 and from the formula $2p_a(D) - 2 = [D \cdot (D + K(V))]$.

(b) *There is a constant* d, *depending only on* c_1, c_2, c_3, *such that* $\Lambda(mdZ)$ *has no base point for* $m \geqq 1$. In fact, there is an integer a such that $3 \leqq a \leqq c_{11} + 3$ and that $s(aZ - Z) = s(aZ)$ (cf. Cor. 2 of Lemma 4). Set $r = t = 1$ in (a). Then we see that $\operatorname{Tr}_Z \Lambda(aZ)$ is complete by Corollary of Lemma 3; moreover, it is of degree $> 2p_a(Z)$ since $\mathfrak{L}(aZ)/\mathfrak{L}(aZ - Z)$ is isomorphic to $\mathfrak{L}((aZ)_\wedge Z)$. When that is so, $\Lambda(aZ)$ has no base point by Lemma 8.

Now let E be an irreducible member of $\Lambda(dZ)$ (cf. Cor. 3 of Lemma 4). By Corollary 2 of Lemma 4, there is an integer b such that $3 \leqq b \leqq c_{11} + 3$ and that $s(bE) = s(bE - E)$. $\operatorname{Tr}_E \Lambda(E)$ has no base point by (b). By (a) and Corollary of Lemma 3, $\operatorname{Tr}_E \Lambda(bE)$ is complete and is of degree $> 2p_a(E)$. Moreover, $\operatorname{Tr}_E \Lambda(E)$ and $\operatorname{Tr}_E \Lambda(bE)$ satisfy the conditions of Lemma 6 by (a) and (b). Therefore, the minimum sum of them is complete, which implies that $\operatorname{Tr}_E \Lambda(bE + E)$ is complete and $s(bE + E) = s(bE)$

(cf. Cor. of Lemma 3). Repeating this, we see that $s(c_{11}E+3E+mE)$ is a constant for all $m \geqq 1$. By (I) and (b), a non-degenerate map h_m of V into a projective space, determined by $\Lambda(m(c_{11}+3)E)$, is a morphism and has no fundamental curve on V. Then it is a projective embedding for large m, which is an easy consequence of the possibility of projective normalization in an algebraic extension of the function field (cf. [7], Chap. IV, Prop. 8). Then h_m is already a projective embedding if

$$m > [c_{11}+3)E\cdot K(V)] + r \geqq 2(c_{11}+3)dc_{12}c_5c_4+4$$

by Lemma 5 and Lemma 7. Our theorem is thereby proved.

Chapter IV.

So far, we have restricted our technique to the use of irreducible curves on the surface. Since the generalized Riemann-Roch theorem is available for such curves, it was easy to see, for instance, whether some linear systems on such curves are free from base points, etc. On the other hand, if we generalize a criterion of ampleness to reducible curves by means of the theory of schemes, we can simplify the latter part of Chapter III to some extent. Let V be a projective variety. Denote by $\mathfrak{O}_V$ the sheaf of local rings on V. If D is a Cartier divisor on V, we mean by the associated subscheme $\mathfrak{D}$ the subscheme of the scheme V (a) whose underlying space is the support of D and (b) whose sheaf $\mathfrak{O}_{\mathfrak{D}}$ is defined at a point x of D to be $\mathfrak{O}_{x,V}/(f)$ for any local equation f of the divisor D, where $\mathfrak{O}_{x,V}$ is the local ring of V at x. We further denote by $\mathfrak{m}_{x,V}$ the maximal ideal of $\mathfrak{O}_{x,V}$. In the following proposition, we discuss a criterion of ampleness on a positive 1-cycle on V. The first half of the proposition has been settled essentially in Lemma 4, and we give only a brief account of the proof for it in the sheaf theoretic terminology.

PROPOSITION. *Let* $T=\sum_{1}^{t} a_iY_i$ *be a positive divisor on a non-singular projective surface* W. *Let* τ *be the subscheme of the scheme* W, *associated to* T, *and* $\mathfrak{M}$ *an invertible sheaf on* τ. *Let* d_i *be the degree of the Cartier divisor class on* Y_i *defined by* $\mathfrak{M}\otimes\mathfrak{O}_{Y_i}$. *If* $d_i > [(K(W)+T')\cdot Y_i]$ *for all divisors* T' *such that* $0 \prec T' \prec T$, *and for all* i *such that* $1 \leqq i \leqq t$, *it follows that* $H^1(\mathfrak{M})=0$. *Moreover, if* $d_i > [(K(W)+T')\cdot Y_i] + 2\max_{x\in Y_i}$ *(multiplicity of* x *on* Y_i*), then* $\mathfrak{M}$ *is ample on* τ.

Proof. We have $\mathfrak{O}_W/\mathfrak{L}(-T) \cong \mathfrak{O}_\tau$, $\mathfrak{O}_W/\mathfrak{L}(-T') \cong \mathfrak{O}_{\tau'}$ from the definitions of $\mathfrak{O}_\tau$, $\mathfrak{O}_{\tau'}$, where τ' is the subscheme of W, associated to T'. Moreover, we have $\mathfrak{L}(-T')/\mathfrak{L}(-T) = \mathfrak{L}(-T') \otimes \mathfrak{O}_{Y_1}$ when we use the same notations as in the proof of Lemma 4. Hence, we have the exact sequence:

$$0 \to \mathfrak{L}(-T') \otimes \mathfrak{O}_{Y_1} \to \mathfrak{O}_T \to \mathfrak{O}_{T'} \to 0.$$

Tensoring the above with $\mathfrak{M}$, we get

$$H^1(\mathfrak{M} \otimes \mathfrak{L}(-T') \otimes \mathfrak{O}_{Y_1}) \to H^1(\mathfrak{M}) \to H^1(\mathfrak{M} \otimes \mathfrak{O}_{T'}) \to 0$$

instead of

$$\begin{aligned} H^1(\mathfrak{L}(A-T')/\mathfrak{L}(A-T)) &\to H^1(\mathfrak{L}(A)/\mathfrak{L}(A-T)) \to \\ &\to H^1(\mathfrak{L}(A)/\mathfrak{L}(A-T')) \to 0 \end{aligned}$$

in the proof of Lemma 4. Computing the degree of the Cartier divisor class on Y_1 determined by $\mathfrak{M} \otimes \mathfrak{L}(-T') \otimes \mathfrak{O}_{Y_1}$, we get $H^1(\mathfrak{M}) = 0$ as in the proof of Lemma 4.

For an invertible sheaf $\mathfrak{L}$ on a complete algebraic scheme W to be ample, it is necessary and sufficient that:

(i) *the sections of* $\mathfrak{L}$ *separate points,*

(*) (ii) *for any point* x *on* W, *if we identify the stalk* $\mathfrak{L}_x$ *with* $\mathfrak{O}_{x,W}$, *the sections of* $\mathfrak{L}$ *which are zero at* x *span* $\mathfrak{m}_{x,W}/\mathfrak{m}_{x,W}^2$.

For every point x in the support $|T|$ of T, let D_x be a positive W-divisor such that $i(D_x \cdot Y_i, x; V)$ gives the multiplicity of x on Y_i for $1 \leqq i \leqq t$. Moreover, for every pair of distinct points x, y in $|T|$, define an invertible sheaf $\mathfrak{M}_{x,y}$ on τ as follows:

$$\begin{aligned} \mathfrak{M}_{x,y} &= \mathfrak{M} \otimes \mathfrak{L}(-D_x) \text{ in } |T| - y - (|T| \cap |D_x| - x), \\ &= \mathfrak{M} \otimes \mathfrak{L}(-D_y) \text{ in } |T| - x - (|T| \cap |D_y| - y), \\ &= \mathfrak{M} \text{ elsewhere.} \end{aligned}$$

Similarly, for every x in $|T|$, define $\mathfrak{M}_{x,x}$ as follows:

$$\begin{aligned} \mathfrak{M}_{x,x} &= \mathfrak{M} \otimes \mathfrak{L}(-2D_x) \text{ in } |T| - (|T| \cap |D_x| - x), \\ &= \mathfrak{M} \text{ elsewhere.} \end{aligned}$$

The first half of the proposition implies $H^1(\mathfrak{M}_{x,y}) = 0$ for all the pairs (x, y), because the degree of the restriction of $\mathfrak{M}_{x,y}$ to Y_i is at least $d_i - 2\max_{z \in Y_i}$ (multiplicity of z on Y_i). Therefore, we have the exact sequence

$$0 \to H^0(\mathfrak{M}_{x,y}) \to H^0(\mathfrak{M}) \to H^0(\mathfrak{M}/\mathfrak{M}_{x,y}) \to H^1(\mathfrak{M}_{x,y}) = 0$$

for all pairs (x, y). Now suppose $x \neq y$. Then $H^0(\mathfrak{M}/\mathfrak{M}_{x,y})$, whose support is the union of two points x and y, contains a section which is 0 at x and a unit at y. Hence $H^0(\mathfrak{M})$ contains a section with the same property. Then the sections of $\mathfrak{M}$ separates points. Let $f = 0$ be a local equation of D_x at x, f being an element of $\mathfrak{O}_{x,W}$. Then f induces an element f' in $\mathfrak{O}_{x,T}$, and we have $(\mathfrak{M}/\mathfrak{M}_{x,x})_x = \mathfrak{O}_{x,T}/f'^2$. But this maps surjectively to the ring $\mathfrak{O}_{x,T}/\mathfrak{m}_{x,T}{}^2$. Hence, if we identify $\mathfrak{M}_x$ to $\mathfrak{O}_{x,T}$, the sections of $\mathfrak{M}$ which are zero at x span $\mathfrak{m}_{x,T}/\mathfrak{m}_{x,T}{}^2$.

Using this proposition, we recover Corollary 1 and Corollary 2 of Lemma 4. Moreover, Corollary 1 can be expressed as $H^1(\mathfrak{L}(mX - U) \otimes \mathfrak{O}_T) = 0$ for $m \geqq c_{10}$, where τ is the subscheme of V, corresponding to a member T of $\Lambda(2c_5X)$. Furthermore, we see that $\mathfrak{L}(mX) \otimes \mathfrak{O}_T$ is ample on τ for $m \geqq c_{10}$ when c_{10} is chosen suitably. Then, by Corollary 2 of Lemma 4, one can find an integer r such that $s(c_{10}T + rT) = s(c_{10}T + rT + T)$ and that $0 \leqq r \leqq c_{11}$. Set $m = 2c_5(c_{10} + r + 1)$. Since $H^1(\mathfrak{L}(mX) \otimes \mathfrak{O}_T) = (0)$ and $s(mX) = s(mX - T)$, it follows that the restriction map $H^0(\mathfrak{L}(mX)) \to H^0(\mathfrak{L}(mX) \otimes \mathfrak{O}_T)$ is surjective. Moreover, $\mathfrak{L}(mX) \otimes \mathfrak{O}_T$ is ample on τ. To prove that $\mathfrak{L}(mX)$ is ample, we again use the criterion (*) cited above. Let x and y be two given points on V. Let T_1 and T_2 be members of $\Lambda(c_5X)$ such that T_1 goes through x and that T_2 goes through y (cf. Lemma 1). When we set $T = T_1 + T_2$, x and y are in the support of τ. Since $\mathfrak{L}(mX) \otimes \mathfrak{O}_T$ is ample, there is a section of this sheaf which is zero at x and not zero at y if $x \neq y$. Lifting this to a section of $\mathfrak{L}(mX)$, the same is true of $\mathfrak{L}(mX)$. If $x = y$, set $T = 2T_1$. The restriction of functions from V to τ induces an isomorphism between $\mathfrak{m}_{x,V}/\mathfrak{m}_{x,V}{}^2$ and $\mathfrak{m}_{x,T}/\mathfrak{m}_{x,T}{}^2$. When we identify the stalk of $\mathfrak{L}(mX) \otimes \mathfrak{O}_T$ at x with $\mathfrak{O}_{x,T}$, the sections of $\mathfrak{L}(mX) \otimes \mathfrak{O}_T$ which vanish at x span $\mathfrak{m}_{x,T}/\mathfrak{m}_{x,T}{}^2$. Lifting these sections, we see that the same is true of $\mathfrak{L}(mX)$ and $\mathfrak{m}_{x,V}/\mathfrak{m}_{x,V}{}^2$. Hence $\mathfrak{L}(mX)$ is ample for $m = 2c_5(c_1 + r + 1)$, where r is a certain integer such that $0 \leqq r \leqq c_{11}$. Since the sum of two ample divisors is ample, we see that $\mathfrak{L}(dX)$ is ample for a suitable d, which depends only on c_1, c_2, c_3. Then we get our Theorem 4 again.

REFERENCES.

[1] S. Abhyankar, "On the valuations centered in a local domain," *American Journal of Mathematics*, vol. 78 (1956), pp. 321-348.

[2] A. Grothendieck, "Éléments de géometrie algébriques, I, II, III," *Publications Mathematiques Institut des Hautes Études Scientifiques*, Paris.

[3] H. Hironaka, "A note on algebraic geometry over ground rings," *Illinois Journal of Mathematics*, vol. 2 (1958), pp. 355-366.

[4] ———, "On the arithmetic genera and the effective genera of algebraic curves," *Memoirs of the College of Science, University of Kyoto*, vol. 30 (1957), pp. 177-195.

[5] K. Kodaira and D. C. Spencer, "On deformations of complex analytic structures, I, II," *Annals of Mathematics*, vol. 67 (1958), pp. 328-466.

[6] M. Kuranishi, "On the locally complete families of complex analytic structures," *Annals of Mathematics*, vol. 75 (1962), pp. 536-577.

[7] S. Lang, *Introduction to Algebraic Geometry*, Interscience Tracts, No. 5, 1958.

[8] T. Matsusaka, "Polarized varieties, fields of moduli and generalized Kummer varieties of polarized Abelian varieties," *American Journal of Mathematics*, vol. 80 (1958), pp. 45-82.

[9] ———, "Q-varieties," *to appear.*

[10] D. Mumford, "Geometric invariant theory," *to appear.*

[11] ———, "Further pathologies in algebraic geometry," *American Journal of Mathematics*, vol. 84 (1962), pp. 642-647.

[12] Y. Nakai, "Non-degenerate divisors on an algebraic surface," *Journal of Science of the Hiroshima University*, vol. 24 (1960), pp. 1-6.

[13] M. Rosenlicht, "Equivalence relations on algebraic curves," *Annals of Mathematics*, vol. 56 (1952), pp. 169-191.

[14] ———, "Generalized Jacobian varieties," *Annals of Mathematics*, vol. 59 (1954), pp. 505-530.

[15] J.-P. Serre, "Faisceaux algébriques coherents," *Annals of Mathematics*, vol. 61 (1955), pp. 197-278.

[16] ———, *Groupes algébriques et corps de classes*, Actualités Scientifiques et Industrielles, No. 1264, Paris.

[17] G. Shimura, "Reduction of algebraic varieties with respect to a discrete valuation of the basic field," *American Journal of Mathematics*, vol. 77 (1955), pp. 134-176.

[18] A. Weil, *Foundations of algebraic geometry*, revised edition, American Mathematical Society Colloquium Publications, No. 29 (1960).

[19] ———, *Variétiés Abéliennes et Courbes algébriques*, Actualités Scientifiques et Industrielles, No. 1064, Paris.

[20] ———, "On the theory of complex multiplications," *Proceedings, International Symposium on Algebraic Number Theory*, Nikko, Japan, 1955.

[21] O. Zariski, "Complete linear systems on normal varieties and a generalization of a lemma of Enriques-Severi," *Annals of Mathematics*, vol. 55 (1952), pp. 552-592.

[22] ———, "Scientific report on the second summer institute, III," *Bulletin of the American Mathematical Society*, vol. 62 (1956), pp. 117-141.

CENTRAL LIBRARY UNIVERSITY

ON TANGENT BUNDLES OF FIBRE SPACES AND QUOTIENT SPACES.

By R. H. Szczarba.*

Introduction. One of the most important invariants associated with a differentiable manifold is its tangent bundle. Indeed, the tangent bundle has been involved, directly or indirectly, in virtually all of the recent progress in differential topology. In spite of this, relatively little has been done toward the determination of the tangent bundle of a given manifold or class of manifolds. (Two exceptions to this statement are Borel and Hirzebruch, [2], Proposition 7.5 and Wu, [12], p. 86.) In this paper, we investigate the tangent bundles of manifolds occurring in differentiable fiber bundles.

The first section of the paper is devoted to the statements of the main theorems. Theorem 1.1 gives information about the tangent bundle of a quotient manifold X/G in terms of a G-equivariant embedding of the manifold X in Euclidean space. The second result, Theorem 1.2 describes the tangent bundle of the total space of a bundle with fiber F and group G in terms of a G-equivariant embedding of F in Euclidean space and Theorem 1.3 is a combination of Theorems 1.1 and 1.2 applying to bundles with fibres X/G.

The next two sections give applications. In Section 2, we apply Theorem 1.1 to quaternionic projective spaces and, in Section 3, to manifolds of constant positive curvature. The remaining four sections give proofs of results stated in the first three sections.

Finally, I would like to express my gratitude to W. S. Massey and L. Auslander for many stimulating and informative conversations during the preparation of this paper.

1. The main theorem. Let ξ be a fibre bundle[1] and suppose a group H acts on E_ξ and B_ξ with $\pi_\xi(hx) = h\pi_\xi(x)$ for $x \in E_\xi$, $h \in H$. Then π_ξ induces a map $\pi_\xi'\colon E_\xi/H \to B_\xi/H$ and, under suitable circumstances, the triple $(E_\xi/H, \pi_\xi', B_\xi/H)$ is again a fiber bundle with fiber F_ξ and group G_ξ (see

---

Received March 14, 1963.

* During the preparation of this paper, the author was partially supported by NSF Grant No. NSF G-18995.

[1] We use the term fiber bundle to mean "espace fibré," as defined in Cartan [3], exposé 6. Our notation is that of Borel-Hirzebruch [2].

Lemma 4.1). Under these circumstances, we say that H acts on the bundle ξ and denote the quotient bundle by ξ/H.

Suppose M is a differentiable manifold. We will denote the tangent bundle of M by $\tau(M)$ and the trivial r-plane bundle over M by $\theta^r = \theta^r(M)$. If ξ is a differentiable fiber bundle, we denote by $\tau_F(\xi)$ the bundle of vectors in $\tau(E_\xi)$ tangent to fibers and by $\tau_\perp(\xi)$ the bundle of vectors in $\tau(E_\xi)$ orthogonal to fibers (in some Riemannian metric).

Clearly $\tau(E_\xi)$ is equivalent to the Whitney sum $\tau_F(\xi) \oplus \tau_\perp(\xi)$ and $\tau_\perp(\xi) = \pi_\xi^{\#}\tau(B_\xi)$, the bundle over E_ξ induced by π_ξ from $\tau(B_\xi)$.

For the remainder of this section, all bundles will be differentiable.

Suppose ξ is a principal bundle with E_ξ compact and let $\phi\colon E_\xi \to R^n$ be an embedding which is equivariant relative to a representation $\alpha\colon G_\xi \to 0(n)$.[2] Giving E_ξ the Riemannian metric induced by ϕ and letting ν_ϕ be the normal bundle of the embedding ϕ, we see that G_ξ acts on both ν_ϕ and $\tau_F(\xi)$ as described above. Now if $\alpha(\xi)$ denotes the n-plane bundle associated with the α-extension of ξ (see [2], p. 477), we have

THEOREM 1.1.

$$\tau(B_\xi) \oplus \tau_F(\xi)/G_\xi \oplus \nu_\phi/G_\xi = \alpha(\xi).$$

For example, suppose $\xi = (S^n, \pi, P_n(R))$ is the standard Z_2-bundle where S^n is the n-sphere and $P_n(R)$ is n-dimensional real projective space. Then, the usual embedding $\phi\colon S^n \to R^{n+1}$ is equivariant relative to the representation $\alpha\colon Z_2 \to 0(n+1)$ which takes the non-zero element of Z_2 into the negative of the identity in $0(n+1)$. Thus by Theorem 1.1,

$$\tau(P_n(R)) \oplus \nu_\phi/Z_2 = \alpha(\xi).$$

Now, $\alpha(\xi)$ is easily seen to be the $(n+1)$-fold Whitney sum of the line bundle $\hat{\xi}$ associated with ξ and ν_ϕ/Z_2 is trivial since ν_ϕ has a Z_2-equivariant cross section. Therefore, we have the well known result[3]

$$\tau(P_n(R)) \oplus \theta^1 = (n+1)\hat{\xi}.$$

The analogous result for complex projective space follows in exactly the same way.

Next, let ξ be a principal bundle and suppose G_ξ acts (on the left) on a manifold F. Let ζ be the associated bundle with fiber F and suppose

---

[2] Mostow [8] has shown that, if E_ξ and G_ξ are compact, such embeddings always exist. See also Palais [9].

[3] This result and the analogue for complex projective space are proved by Milnor in his notes on characteristic classes pp. 10-12 and 74-75. See also Atiyah [1], Lemma 4.5.

$\psi\colon F\to R^n$ is embedding which is equivariant relative to a representation $\beta\colon G_\xi\to 0(n)$. Now, since $E_\zeta = E_\xi\times_{G_\xi}F$ and $E_{\beta(\xi)} = E_\xi\times_{G_\xi}R^n$, ψ induces an embedding $\phi\colon E_\xi\to E_{\beta(\xi)}$.

THEOREM 1.2.

$$\tau_F(\zeta)\oplus\nu_\phi = \pi_\zeta^{\#}\beta(\xi).$$

To illustrate, suppose ξ is a principal $0(n)$-bundle, $\hat{\xi}$ the associated vector bundle, and ζ the associated sphere bundle. The usual embedding $\psi\colon S^{n-1}\subset R^n$ is $0(n)$-equivariant and, since ν_ψ has an $0(n)$-equivariant cross section, the normal bundle of the induced embedding $\phi\colon E_\zeta\to E_{\hat{\xi}}$ is easily seen to be trivial. Thus,

$$\tau_F(\zeta)\oplus\theta^1 = \pi_\zeta^{\#}\hat{\xi}$$

where θ^1 is the trivial line bundle over E_ζ (see Wu [12], p. 86).

Now suppose ξ and η are principal bundles and suppose G_ξ acts on E_η with $(gx)g' = g(xg')$ for $g\in G_\xi$, $g'\in G_\eta$, and $x\in E_\eta$. Then, G_ξ acts on B_η and we can form the bundles ξ_1 with fiber E_η, ξ_2 with fiber B_η associated with ξ. Furthermore, the map $1\times\pi\colon E_\xi\times E\to E_\xi\times B$ induces maps

$$\pi_1\colon E_\xi\times E_\eta \longrightarrow E_\xi\times_{G_\xi}B_\eta = E_{\xi_2},$$
$$\pi_2\colon E_\xi\times_{G_\xi}E_\eta = E_{\xi_1}\to E_\xi\times_{G_\xi}B_\eta = B_{\xi_2}$$

and the triples $\zeta_1 = (E_\xi\times E_\eta, \pi_1, E_{\xi_2})$, $\zeta_2 = (E_{\xi_1}, \pi_2, E_{\xi_2})$ are principal bundles. (ζ_1 is a principal $G_\xi\times G_\eta$ bundle and ζ_2 a principal G_η bundle. See Lemma 4.2 below.)

Let $G_\xi\times G_\eta$ act on E_η by $(g_1,g_2)x = g_1xg_2^{-1}$ and let $\psi\colon E_\eta\to R^n$ be an embedding equivariant relative to a representation $\gamma\colon G_\xi\times G_\eta\to 0(n)$. Then $1\times\psi\colon E_\xi\times E_\eta\to E_\xi\times R^n$ induces an embedding $\phi\colon E_{\xi_1}\to E_{\alpha(\xi)}$ where $\alpha\colon G_\xi\to 0(n)$ is the restriction of γ to $G_\xi\times 1$. (The diagram of Section 5.1 should help clarify the situation here.)

THEOREM 1.3. *With notation as above,*

$$\tau_F(\xi_2)\oplus\tau_F(\zeta_2)/G_\eta\oplus\nu_\phi/G_\eta = \gamma(\zeta_1).$$

For example, let ξ be a principal $0(n+1)$-bundle and $\eta = (S^n, \pi, P_n(R))$ the standard principal Z_2-bundle. Denote by ξ_1 the S^n-bundle associated with ξ, by ξ_2 the $P_n(R)$-bundle associated with ξ, by ζ_1 the principal $Z_2\times 0(n+1)$-bundle $\pi_1\colon E_\xi\times S^n\to E_{\xi_2}$, and by ζ_2 the principal Z_2-bundle $\pi_2\colon E_{\xi_1}\to E_{\xi_2}$. If $\psi\colon S^n\to R^{n+1}$ is the usual embedding and $\gamma\colon Z_2\times 0(n+1)\to 0(n+1)$ the multiplication map (identifying Z_2 with the center of $0(n+1)$), it is easily seen that ψ is equivariant relative to the representation γ so, by Theorem 1.3, we have $\tau_F(\xi_2) \oplus \nu_\phi/Z_2 = \gamma(\zeta_1)$.

Now, if $\gamma_1 : Z_2 \times 0(n+1) \to Z_2 = 0(1)$ and $\gamma_2 : Z_2 \times 0(n+1) \to 0(n+1)$ are the projection maps, one sees easily that $\gamma(\zeta_1) = \gamma_1(\zeta_1) \otimes \gamma_2(\zeta_1)$, $\gamma_1(\zeta_1) = \hat{\zeta}_2$, and $\gamma_2(\zeta_1) = \pi_{\xi_2}{}^*\hat{\xi}$. (We use $\hat{\zeta}_2$ to denote the line bundle associated with ζ_2 and $\hat{\xi}$ the $(n+1)$-plane bundle associated with ξ.) Furthermore, since the embedding $\psi : S^n \to R^{n+1}$ has a $Z_2 \times 0(n+1)$ equivariant normal field, it follows that ν_ϕ/Z_2 is trivial and we have

$$\tau_F(\xi_2) \oplus \theta^1 = \zeta_2 \otimes \pi_{\xi_2}{}^*\hat{\xi}.$$

As a consequence, we have the formula of [2], p. 517 (see [13], Lemma 5.1).

$$\omega(\tau_F(\xi_2)) = \sum_{i+j=n+1} (1 + \omega_1(\hat{\zeta}))^i \pi_{\xi_2}{}^*\omega_j(\hat{\xi}).$$

In exactly the same way, we can treat bundles with fiber $P_n(C)$ and group $U(n+1)$ and bundles with lens spaces as fiber and suitably restricted groups.

2. The tangent bundle of quaternionic projective space. Let α: $Sp(1) \to SO(4n+4)$ be the composite.

$$Sp(1) \to Sp(1) \times \cdots \times Sp(1) \subset Sp(n+1) \subset SO(4n+4)$$

where the first map is the diagonal map. Then $Sp(1)$ acts on $S^{4n+3} \subset R^{4n+4}$ via α and defines a principal $Sp(1)$-bundle ξ with $E_\xi = S^{4n+3}$ and $B_\xi = P_n(H)$, the n-dimensional quaternionic projective space. Then, by Theorem 1.1, we have

$$\tau(P_n(H)) \oplus \tau_F(\xi)/Sp(1) \oplus \nu/Sp(1) = \alpha(\xi)$$

where ν is the normal bundle to the embedding $S^{4n+3} \subset R^{4n+4}$. In fact, $\alpha(\xi)$ is the $(n+1)$-fold Whitney sum $(n+1)\hat{\xi}$ where $\hat{\xi}$ is the 4-plane bundle associated with ξ and $\nu/Sp(1)$ is trivial. Further, if β: $Sp(1) \to SO(3)$ is the 2-fold covering (see Steenrod [11], p. 115), we prove in Section 6 that $\tau_F(\xi)/Sp(1) = \beta(\xi)$. Thus we have[4] (compare Lemma 4.5 of [1]).

THEOREM 2.1. $\tau(P_n(H)) \oplus \beta(\xi) \oplus \theta^1 = (n+1)\hat{\xi}$.

For any vector bundle ζ, we denote the Stiefel-Whitney class by $w(\zeta) = \sum w_i(\zeta)$ and the Pontrjagin class by $p(\zeta) = \sum p_i(\zeta)$. If M is a manifold, $w(M) = w(\tau(M))$ and $p(M) = p(\tau(M))$.

---

[4] In [7] the author and W. C. Hsiang prove an analogue of Theorem 2.1 for real, complex, and quaternionic Grassmann manifolds.

To compute the characteristic classes of $P_n(H)$, we need the following.

LEMMA 2.2. *Let v be a generator of $H^4(P_n(H);Z)$ and $u \in H^4(P_n(H);Z_2)$ its reduction* mod 2. *Then*

$$w(\beta(\xi))=1, \qquad p(\beta(\xi))=1+4v,$$
$$w(\hat{\xi})=1+u, \qquad p(\hat{\xi})=(1+v)^2.$$

We will prove this lemma in Section 7. As an immediate consequence, we have (see Hirzebruch [5] and Borel and Hirzebruch [2], pp. 517-520)

COROLLARY 2.3. *Let $v \in H^4(P_n(H);Z)$ be a generator and $u \in H^4(P_n(H);Z)$ its reduction* mod 2. *Then*

$$w(P_n(H))=(1+u)^{n+1},$$
$$p(P_n(H))=(1+v)^{2n+2}(1+4v)^{-1}.$$

3. Manifolds of constant positive curvature. Let M be a compact orientable Riemannian n-manifold of constant positive curvature. As is well known (see Hopf [6]), M is finitely covered by the n-sphere S^n where the group of covering transformations G acts on S^n via a representation α: $G \to O(n+1)$. Let $\xi=(S^n,\pi,M)$ denote the covering and ν the normal bundle of the embedding $S^n \subset R^{n+1}$. Since ν clearly has a G-equivariant cross section, ν/G is trivial and we have

THEOREM 3.1. $\tau(M) \oplus \theta^1=\alpha(\xi)$.

We now consider some special cases.

Let Z_m be the cyclic group of order m with generator g. For any integer q relatively prime to m, we define a representation $\alpha(q): Z_m \to SO(2)$ by

$$\alpha(q)(g)=\begin{vmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{vmatrix}$$

where $\gamma=2\pi q/m$. If $q_1,\cdots,q_n$ are relatively prime to m, we let Z_m act on $S^{2n+1} \subset R^{2n+2}$ via the direct sum $\alpha(1)\oplus\alpha(q_1)\oplus\cdots\oplus\alpha(q_n)$. This action defines a principal Z_m bundle ξ with $E_\xi=S^{2n+1}$ and B_ξ is the lens space $L(m;q_1,\cdots,q_n)$.

COROLLARY 3.2. *Let $L=L(m;q_1,\cdots,q_n)$ and ξ be the principal Z_m-bundle over L. Then*

$$\tau(L)\oplus\theta^1=\xi_0\oplus\xi_1\oplus\cdots\oplus\xi_n$$

where ξ_0 is the 2-plane bundle associated with ξ and ξ_j is the 2-plane bundle *associated with the* $\alpha(q_j)$*-extension of* ξ. *Furthermore, if* v *is a generator of* $H^2(L;Z)$ *and* u *its reduction* mod 2, *then*

$$p(L) = (1+v^2)\prod_{j=1}^{n}(1+q_j^2v^2),$$

$$w(L) = (1+u)\prod_{j=1}^{n}(1+q_ju).$$

The first part of the corollary is an immediate consequence of Theorem 3.1 whereas the expressions for the characteristic classes follow easily from the fact that the $\alpha(q)$-extension of ξ has classifying map

$$L \subset B_{Z_m} \xrightarrow{\lambda} B_{SO(2)}.$$

Here L is considered the $(2n+1)$-skeleton of B_{Z_m} and λ is induced by the homomorphism $\alpha(q): Z_m \to SO(2)$.

Let H_m denote the generalized quaternion group with generators a and b and relations $aba = b$ and $a^r = b^2$ where $r = 2^{m-1}$. For any odd integer q, let $\beta(q): H_m \to SO(4)$ be the representation defined by

$$\beta(q)(a) = \begin{vmatrix} \cos\gamma & \sin\gamma & \multicolumn{2}{c}{0} \\ -\sin\gamma & \cos\gamma & & \\ & & \cos\gamma & -\sin\gamma \\ \multicolumn{2}{c}{0} & \sin\gamma & \cos\gamma \end{vmatrix},$$

$$\beta(q)(b) = \begin{vmatrix} 0 & I \\ -I & 0 \end{vmatrix}$$

where I is the 2×2 identity matrix and $\gamma = q\pi/2^{m-1}$.

If $q_1, \cdots, q_n$ is a sequence of odd integers, we let H_m act on $S^{4n+3} \subset R^{4n+4}$ via the direct sum $\beta(1) \oplus \beta(q_1) \oplus \cdots \oplus \beta(q_n)$. The action defines a principal H_m-bundle ξ with $E_\xi = S^{4n+3}$ and $B_\xi = N(m; q_1, \cdots, q_n)$.

COROLLARY 3.3. *Let* $N = N(m; q_1, \cdots, q_n)$ *and* ξ *be the principal* H_m*-bundle over* N. *Then*

$$\tau(N) \oplus \theta^1 = \xi_0 \oplus \xi_1 \oplus \cdots \oplus \xi_n$$

where $\xi_0 = \beta(1)(\xi)$ *and* $\xi_j = \beta(q_j)(\xi)$.

Furthermore, if u *is the non-zero element of* $H^4(N;Z_2) \approx Z_2$,

$$w(N) = (1+u)^{n+1}. \tag{3.1}$$

The first part of Corollary 3.3 follows immediately from Theorem 3.1. We will prove equation (3.1) in Section 7.

Remark. Note that, since the homomorphism $\beta(q): H(m) \to SO(4)$ factors through $Sp(1) \subset SO(4)$, the group of the bundle $\tau(N) \oplus \theta^1$ admits a reduction to the sympletic group. Thus we can say that N has a generalized almost quaternionic structure.

We return now to the general case. Let M be a compact orientable manifold of constant positive curvature and let S be a p-sylow subgroup of the fundamental group G of M. Then S is cyclic for $p > 2$ and either cyclic or a generalized quaternion group for $p = 2$ (see P. Smith, [10]). Considering M as S^n/G, we see that the inclusion may $S \subset G$ induces a map of S^n/S onto S^n/G and S^n/S is either a real projective space, a a lens space, or an $N(m, q_1, \cdots, q_n)$ defined above. (Here $n = 4r + 3$ and 2^{m+1} is the order of S.) Furthermore, the map of S^n/S onto M induces a monomorphism on mod p cohomology (see, for example, Cartan-Eilenberg, [4], p. 259.) Thus, since we know the characteristic classes of S^n/S, we can compute the Stiefel-Whitney classes of M as well as the mod p components of the Pontrjagin classes for any prime p.

4. Two preliminary lemmas.

LEMMA 4.1. *Let ξ be a principal bundle with E_ξ compact and suppose a group H acts (on the left) on E_ξ such that*

(4.1) $(hx)g = h(xg)$ *for* $h \in H$, $x \in E_\xi$, *and* $g \in G_\xi$,

(4.2) *the induced action of H on B_ξ is without fixed points,*

(4.3) *the spaces E_ξ/H and B_ξ/H are Hausdorff.*

Then the triple $(E_\xi/H, \pi_\xi', B_\xi/H)$ is a principal G_ξ-bundle where π_ξ' is induced by π_ξ.

Proof. All we need to show is that, for any $x, y \in E_\xi/H$ with $\pi_\xi' x = \pi_\xi' y$, there is a unique $g \in G_\xi$ with $xg = y$ and that the resulting map of

$$A' = \{(x, y) \in E_\xi/H \times E_\xi/H \ni \pi_\xi' x = \pi_\xi' y\}$$

into G_ξ is continuous.

The fact that there is a unique $g \in G_\xi$ with $xg = y$ for $(x, y) \in A'$ follows from the corresponding property for ξ. To see that the resulting map

$f'\colon A' \to G_\xi$ is continuous, we let $A = \{(x,y) \in E_\xi \times E_\xi \ni \pi_\xi x = \pi_\xi y\}$ and let $f\colon A \to G_\xi$ be the map with $xf(x,y) = y$. Now H acts on A by $h(x,y) = (hx,hy)$ and $f(hx,hy) = f(x,y)$ (by (4.1)) so f induces a map $f_1\colon A/H \to G_\xi$. Furthermore, the identity map of $E_\xi \times E_\xi$ onto itself induces a map $\gamma\colon A/H \to A'$ which is clearly continuous, onto, and, by (4.2), one-one. In fact, since E_ξ is compact and A closed in $E_\xi \times E_\xi$, A/H is compact so γ is a homomorphism. Letting $f'\colon A' \to G_\xi$ be the composite $f_1\gamma^{-1}$, we see that $xf'(x,y) = y$ for $(x,y) \in A'$. This completes the proof of Lemma 4.1.

Now, let ξ and H be as in the lemma and suppose G_ξ acts on F. Then, we can form the F-bundle associated with both ξ and ξ/H. In fact, H acts on $E_\xi \times_{G_\xi} F$ (on the first factor) and one sees easily that $(E_\xi \times_{G_\xi} F)/H = (E_\xi/H) \times_{G_\xi} F$. Thus the F-bundle associated with ξ/H is the quotient by H of the F-bundle associated with ξ.

LEMMA 4.2. *Let ξ and ζ be principal bundles and suppose G_ξ acts on E_ζ with $g(xh) = g(xh)$, $g \in G_\xi$, $x \in E_\zeta$, and $h \in G_\zeta$. Then G_ξ acts on B_ζ and the triple $(E_\xi \times_{G_\xi} E_\xi, \pi, E_\xi \times_{G_\xi} B_\xi)$ is a principal G_ξ-bundle where π is induced by the map $1 \times \pi_\zeta\colon E_\xi \times E \to E_\xi \times B_\zeta$.*

Proof. Let

$$\begin{aligned} A &= \{(z,z') \in E_\xi \times_{G_\xi} E_\zeta \times E_\xi \times_{G_\xi} E_\zeta \ni \pi z = \pi z'\}, \\ A_1 &= \{(x,y) \in E_\xi \times E_\xi \ni \pi_\xi x = \pi_\xi y\}, \\ A_2 &= \{(x,y) \in E_\zeta \times E_\zeta \ni \pi_\zeta x = \pi_\zeta y\}, \end{aligned}$$

and let $f_1\colon A_1 \to G_\xi$, $f_2\colon A_2 \to G_\zeta$ be the maps with $xf_i(x,y) = y$ for $(x,y) \in A_i$, $i = 1, 2$. Then we define $f\colon A \to G_\zeta$ by the formula $f(z,z') = f_2(f_1(x,x')y,y')$ where (z,z') in A is represented by (x,y,x',y') in $E_\xi \times E_\zeta \times E_\xi \times E_\zeta$. The verification that f is well defined and has the required properties is left to the reader.

5. The proofs of the main theorems.

Proof of Theorem 1.1. Since $\tau(B_\xi) = \tau_\perp(\xi)/G_\xi$ and $\tau_\perp(\xi) \oplus \tau_F(\xi) = \tau(E_\xi)$, it suffices to show that $(\tau(E_\xi) \oplus \nu_\phi)/G_\xi = \alpha(\xi)$. Now, $\tau(E_\xi) \oplus \nu_\phi$ is trivial and can be considered as the bundle over E_ξ induced by ϕ from the trivial n-plane bundle over R^n. Thus, we have an equivalence $F\colon E \to E_\xi \times R^n$ where E is the total space of $\tau(E_\xi) \oplus \nu_\phi$. In fact, this equivalence is G_ξ-equivariant so induces an equivalence between E/G_ξ and $E_\xi \times_{G_\xi} R^n$. However, E/G_ξ is the total space of $(\tau(E_\xi) \oplus \nu_\phi)/G_\xi$ and $E_\xi \times_{G_\xi} R^n = E_{\alpha(\xi)}$ so Theorem 1.1 is proved.

Proof of Theorem 1.2. We first prove the following lemma.

LEMMA 5.1. *Let* η *be a differentiable* n*-plane bundle. Then* $\tau_F(\eta) = \pi_\eta^{\#}\eta$.

Proof. Suppose η is associated with the $0(n)$-bundle ζ and $\tau = \tau(R^n)$. Then the total space of $\tau_F(\eta)$ is simply $E_\zeta \times_{0(n)} E_\tau$. (See [2], p. 478). The lemma now follows easily from the fact that $E_\tau = R^n \times R^n$.

To prove Theorem 1.2, we first note that the bundle $\tau_F(\zeta) \oplus \nu_\phi$ over E_ζ is induced by $\phi\colon E_\zeta \to E_{\beta(\xi)}$ from the bundle $\tau_F(\beta(\xi))$. Now, by Lemma 5.1, $\tau_F(\beta(\xi))$ is induced by $\pi_{\beta(\xi)}\colon E_{\beta(\xi)} \to B_\xi$ from $\beta(\xi)$ and, since $\pi_{\beta(\xi)}\phi = \pi_\xi$, the theorem is proved.

Proof of Theorem 1.3. The following diagram should help the reader keep track of the bundles involved in this proof.

$$\begin{tikzcd}
E_\xi \times E_\eta \arrow[r, "1 \times \psi"] \arrow[d, "p_1"'] & E_\xi \times R^n \arrow[d, "p_2"] \\
E_\xi \times_{G_\xi} E_\eta \arrow[r, "\phi"] \arrow[d, "\pi_2"'] \arrow[dr, "\pi_{\xi_1}"] & E_\xi \times_{G_\xi} R^n \arrow[d, "\pi_{\alpha(\xi)}"] \\
E_\xi \times_{G_\xi} B_\eta \arrow[r, "\pi_{\xi_2}"'] & B_\xi
\end{tikzcd}$$

Note, first of all that the map $\pi_2\colon E_\xi \times_{G_\xi} E_\eta \to E_\xi \times_{G_\xi} B_\eta$ induces a bundle epimorphism $\pi_2\colon \tau_F(\xi_1) \to \tau_F(\xi_2)$ with kernel $\tau_F(\zeta_2)$. Choosing a Riemannian metric on $E_\xi \times_{G_\xi} E_\eta$, we see that $\tau_F(\xi_1)$ is equivalent to $\tau_F(\zeta_2) \oplus \pi_2^{\#}\tau_F(\zeta_2)$ so that $\tau_F(\zeta_1)/G = [\tau_F(\zeta_2) \oplus \pi_2^{\#}\tau_F(\xi_2)]/G_\eta$. In fact, if we choose the Riemannian metric on $E_\xi \times_{G_\xi} E_\eta$ to be invariant under the action of G_η, we have

$$[\tau_F(\zeta_2) \oplus \pi_2^{\#}\tau_F(\xi_2)]/G_\eta = \tau_F(\zeta_2)/G_\eta \oplus \pi_2^{\#}\tau_F(\xi_2)]/G_\eta.$$

Now $\pi_2^{\#}\tau_F(\xi_2)/G_\eta$ is easily seen to be equivalent to $\tau_F(\xi_2)$ so, to prove the theorem, we need only show that $[\tau_F(\xi_1) \oplus \nu_\phi]/G_\eta = \gamma(\xi)$.

Let $p\colon E_\xi \to E_\eta$ be the projection map. It is easily seen that $\tau_F(\xi_1) = p^{\#}\tau(E_\eta)/G_\xi$ and that $\nu_\phi = p_{\#}\nu_\psi/G_\xi$ (see [2], p. 478) so that

$$\begin{aligned}[\tau_F(\xi_1) \oplus \nu_\phi]/G_\eta &= [p^{\#}\tau(E_\eta)/G_\xi \oplus p_{\#}\nu\psi/G_\xi]/G_\eta \\ &= p^{\#}[\tau(E_\eta) \oplus \nu\psi]/G_\xi \times G_\eta.\end{aligned}$$

Now $\tau(E_\eta) \oplus \nu\psi$ is trivial so the total space of $[\tau_F(\xi_1) \oplus \nu_\phi]/G_\eta$ is $E_\xi \times E_\eta \times R^n/G_\xi \times G_\eta$ and, since $\psi\colon E_\eta \to R^n$ is equivariant, the action of $G_\xi \times G_\eta$ on $E_\xi \times E_\eta \times R^n$ is exactly the action which defines $\gamma(\zeta_1)$. This completes the proof of the theorem.

6. The proof of Theorem 2. 1. We will need the following lemma.

LEMMA 6. 1. *Suppose* ξ *and* η *are principal bundles with* $E_\eta = B_\xi$ *and suppose* G_η *acts (on the right) on* ξ *with* $(g'x)g = g'(xg)$ *for* $g \in G_\eta$, $x \in E_\xi$, *and* $g' \in G_\xi$. *Let* $A = \{(x,y) \in E_\xi \times E_\xi$ *with* $\pi_\xi x = \pi_\xi y\}$, $f\colon A \to G_\xi$ *the map with* $x = f(x,y)y$, *and suppose* $s\colon B_\xi \to E_\xi$ *is a cross section with the property that* $f(s(xg^{-1})g, s(x))$ *is independent of* x *for all* $g \in G_\eta$. *Then the map* $\beta\colon G_\eta \to G_\xi$ *defined by* $\beta(g) = f(s(xg^{-1})g, s(x))$ *is a homomorphism and* ξ/G_η *is equivalent to the* β*-extension of* η.

Proof. First of all,

$$\begin{aligned}\beta(g_1)\beta(g_2) &= f(s(xg_1^{-1})g_1, s(x))f(s(xg_2^{-1})g_2, s(x)) \\ &= f(s(xg_2^{-1}g_1^{-1})g_1, s(xg_2^{-1}))f(s(xg_2^{-1})g_2, s(x))\end{aligned}$$

since $f(s(xg^{-1})g, s(x))$ is independent of x. But then $(\beta g_1)\beta(g_2))s(x) = s(xg_2^{-1}g_1^{-1})g_1g_2$ so

$$\begin{aligned}\beta(g_1)\beta(g_2) &= f(s(xg_2^{-1}g_1^{-1})g_1g_2, s(x)) \\ &= \beta(g_1g_2)\end{aligned}$$

and β is a homomorphism.

Now define $\psi\colon E_\xi \to E_\eta \times G_\xi$ by $\psi(x) = (\pi_\xi(x), f(s\pi_\xi(x), x))$. Then, for $g \in G_\eta$,

$$\begin{aligned}\psi(xg) &= (\pi_\xi(xg), f(s\pi_\xi(xg), xg) \\ &= (\pi_\xi(x)g, f(s(\pi_\xi(x)g), xg)) \\ &= (zg, f(s(zg)g^{-1}, x)) \\ &= (zg, f(s(zg)g^{-1}, s(z))f(s(z), x)) \\ &= (zg, \beta(g)^{-1}f(s(z), x))\end{aligned}$$

where $z = \pi_\xi(x)$. Thus ψ defines a bundle map $\psi_0\colon E_\xi/G_\eta \to E_\eta \times_{G_\eta} G_\xi$ (where G_η acts on G_ξ via β) which is easily seen to be an equivalence.

Theorem 2.1 is an immediate consequence of the following.

COROLLARY 6.2. *Let* η *be a principal bundle,* ξ *the principal* $0(m)$*-bundle associated with* $\tau_F(\eta)$, *and* $\mathrm{Ad}\colon G_\eta \to 0(m)$ *the adjoint representation. Then* ξ/G_η *is equivalent to the* Ad*-extension of* η.

Proof. Let $T(G_\eta)_x$ denote the tangent space to G_η at $x \in G_\eta$, R_g: $T(G_\eta)_x \to T(G_\eta)_{xg}$ the map induced by right translation, and L_g: $T(G_\eta)_x \to T(G_\eta)_{gx}$ the map induced by left translation. Then

$$\mathrm{Ad}(g) = R_g L_{g^{-1}}\colon T(G_\eta)_e \to T(G_\eta)_e.$$

Let $\mu\colon E_\eta \times G_\eta \to E_\eta$ be the principal map and $d\mu_x\colon T(G_\eta)_e \to T(E_\eta)_x$ the map induced by $\mu_x\colon G_\eta \to E_\eta$, $\mu_x(g) = \mu(x, g)$. Clearly $d\mu_x$ takes $T(G_\eta)_e$ isomorphically onto the fiber of $\tau_F(\xi)$ containing x. We define a cross section s in ξ by

$$s(x) = [d\mu_x(v_1), \cdots, d\mu_x(v_m)]$$

where $v_1, \cdots, v_m$ is a base for $T(G_\eta)_e$ and $[d\mu_x(v_1), \cdots, d\mu_x(v_m)]$ denotes the frame determined by $d\mu_x(v_1), \cdots, d\mu_x(v_m)$.

Now,

$$\begin{aligned} s(xg^{-1})g &= [R_g d\mu_{xg^{-1}}(v_1), \cdots, R_g d\mu_{xg^{-1}}(v_m)] \\ &= [R_g L_{g^{-1}} d\mu_x(v_1), \cdots, R_g L_{g^{-1}} d\mu_x(v_m)] \end{aligned}$$

since $\mu(\mu(x, g^{-1}), g') = \mu(x, g^{-1}g')$. Therefore, the map $\beta\colon G_\eta \to 0(m)$ defined in Lemma 6.1 can be identified with the adjoint representation and the corollary is proved.

7. The proofs of Lemma 2.2 and Corollary 3.3.

Let

$$w = \sum w_j \in H^*(B_{O(n)}; Z_2)$$

be the universal Stiefel-Whitney class and $p = \sum p_j \in H^*(B_{SO(n)}; Z)$ the universal Pontrjagin class. We will need the following three lemmas.

LEMMA 7.1. *Let* $j\colon Sp(1) \subset SO(4)$ *be the standard inclusion and* $\lambda(j)\colon B_{Sp(1)} \to B_{SO(4)}$ *the induced map. Then*

$$\lambda(j)^* p = (1 + v)^2, \tag{7.1}$$

$$\lambda(j)^* w = 1 + u, \tag{7.2}$$

where v is a generator of $H^4(B_{Sp(1)}; Z)$ and u its reduction mod 2.

Proof. Let $S^1 \subset Sp(1)$ be unit quaternions of the form $a + bi$. Then

S^1 is a maximal torus. If $S^1 \times S^1$ is the usual maximal torus in $SO(4)$ and $d\colon S^1 \to S^1 \times S^1$ the diagonal map, then the diagram

$$\begin{tikzcd}
S^1 \arrow[r, "d"] \arrow[d] & S^1 \times S^1 \arrow[d] \\
Sp(1) \arrow[r, "j"'] & SO(4)
\end{tikzcd}$$

is commutative. Passing to classifying spaces, we obtain a corresponding diagram in cohomology. Equation (7.1) is an immediate consequence of this diagram (see [2], p. 487) and equation (7.2) follows from the fact that p_i reduced mod 2 is w_{2i}^2.

LEMMA 7.2. *Let* $\lambda(\beta) : B_{Sp(1)} \to B_{SO(3)}$ *be induced by the two fold covering* $\beta\colon Sp(1) \to SO(3)$. *Then*

$$\lambda(\beta)^* p = 1 + 4u$$

where u *is a generator of* $H^4(B_{Sp(1)}; Z)$.

Proof. The short exact sequence

$$1 \to Z_2 \to Sp(1) \longrightarrow SO(3) \to 1$$

defines a fiber map $\lambda(\beta) : B_{Sp(1)} \to B_{SO(3)}$ with fiber B_{Z_2}. If (E_r, d_r) denotes the integral cohohomolgy spectral sequence of this fiber space, it is not difficult to show that $E_\infty^{0,4} \approx E_\infty^{2,2} \approx Z_2$, $E_\infty^{4,0} \approx Z$, and $E_\infty^{1,3} \approx E_\infty^{1,3} = 0$. From this, it is immediate that $H^4(B_{Sp(1)}; Z)/\lambda(\beta)^* H^4(B_{SO(3)}; Z) \approx Z_4$ and the lemma follows.

LEMMA 7.3. *Let* $\lambda(i) : B_{H_m} \to B_{Sp(1)}$ *be induced by the inclusion* i: $H_m \subset Sp(1)$. *Then* $\lambda(i)^* \colon H^*(B_{Sp(1)}; Z_2) \to H^*(B_{H_m}; Z_2)$ *is an isomorphism into.*

Proof. In fact, $\lambda(i) : B_{H_m} \to B_{Sp(1)}$ is a fiber map with fiber $Su(1)/H_m$. Using the fact that $H^3(B_{H_m}; Z_2)$ is non zero (see [4], p. 254), we see that the mod 2 cohomology spectral sequence is trivial. This proves the lemma.

Now, Lemma 2.2 follows immediately from Lemmas 7.1 and 7.2. To prove equation (3.1) we notice that the map of H_m into $SO(4)$ factors into the composite $H_m \xrightarrow{i} Sp(1) \xrightarrow{j} SO(4)$. The result now follows from Lemmas 7.1 and 7.3.

YALE UNIVERSITY,
NEW HAVEN, CONN.

REFERENCES.

[1] M. F. Atiyah, "Thom complexes," *Proceedings of the London Mathematical Society* (3), vol. 11 (1961), pp. 291-310.
[2] A. Borel and F. Hirzebruch, "Characteristic classes of homogeneous spaces I," *American Journal of Mathematics*, vol. 80 (1958), pp. 458-538.
[3] H. Cartan, *Séminaire Henri Cartan 1949-50*, Paris, 1956.
[4] H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton University Press.
[5] F. Hirzebruch, "Über die Quaternionalen Projektiven Räume," *Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Mathematik-Naturwissenschaftliche Klasse*, 1954, 301-312 (1954).
[6] H. Hopf, "Zum Clifford-Kleinschen Räum problem," *Mathematische Annalen*, vol. 95 (1926), pp. 313-339.
[7] W. C. Hsiang and R. H. Szczarba, "On the tangent bundle of a Grassman manifold," *American Journal of Mathematics*, vol. 86 (1964), pp. 698-704.
[8] G. D. Mostow, "Equivariant embeddings in Euclidean space," *Annals of Mathematics*, vol. 65 (1957), pp. 432-446.
[9] R. S. Palais, "Imbedding of compact, differentiable transformation groups in orthogonal representations," *Journal of Mathematics and Mechanics*, vol. 6 (1957), pp. 673-678.
[10] P. Smith, "Permutable periodic transformations," *Proceedings of the National Academy of Sciences, U. S. A.*, vol 30 (1944), pp. 105-108.
[11] N. Steenrod, *The Topology of Fibre Bundles*, Princeton University Press, 1951.
[12] W. Wu, "Sur les Classes Caracteristiques des Structures Fibres Spheriques," *Actualites Scientifiques et Industrielles*, no. 1183, Hermann, Paris, 1952.
[13] Emery Thomas, "On the tensor product of n-plane bundles," *Archiv der Mathematik*, vol. 10 (1959), pp. 174-179.

ON THE TANGENT BUNDLE OF A GRASSMAN MANIFOLD.

By W. C. Hsiang and R. H. Szczarba.[1]

Introduction. The purpose of this paper is to investigate the tangent bundles of the real, complex, and quaternionic Grassman manifolds. The techniques used are those of [7] and the principal result takes the form of an equation relating the tangent bundle of the Grassman manifold to a canonical vector bundle over the Grassman manifold. As a consequence of this equation, we are able to show that the quaternionic Grassman manifolds admit no generalized almost complex structure. This generalizes the known result for quaternionic projective spaces (see Hirzebruch [4] and Massey [6]).

The paper is divided into four sections. The first section contains the statements of the principal results, the second and third sections contain the proofs of these results and the final section is an appendix in which we derive an expression for the Pontrjagin class of a quaternionic tensor product.

The authors would like to express their gratitude to W. S. Massey and G. B. Seligman for several profitable conversations during the preparation of this paper.

1. Statement of results. Let K denote either the field R of real numbers, the field C of complex numbers, or the field H of quaternions and K^n the n-dimensional vector space over K. If ζ is a right K-vector bundle over X, we denote by ζ^* the conjugate bundle (a left K-vector bundle) and by ζ_R the real vector bundle determined by ζ. If η is a left K-vector bundle over X, we can define the tensor product $\zeta \otimes_K \eta$. In case K is the real or complex field, $\zeta \otimes_K \eta$ is a K-vector bundle while if $K = H$, $\zeta \otimes_K \eta$ has a real structure only. We agree, in all thre cases, to treat $\zeta \otimes_K \eta$ only as a real vector bundle ignoring any additional structure.

Let $G_{n-k,k} = G_{n-k,k}(K)$ denote the Grassman manifold of k-planes in K^n and $\xi = \xi_{n-k,k}$ the canonical K-vector bundle over $G_{n-k,k}$. (The total space E_ξ consists of pairs (x, v) where x is an element of $G_{n-k,k}$ and v is a vector in x.)

Received March 14, 1963.

[1] During the preparation of this paper, both authors were partially supported by NSF grant number NSF G18995.

THEOREM 1.1.[2] *Let* $\tau(G_{n-k,k})$ *denote the real tangent bundle of* $G_{n-k,k}$. *Then*

$$\tau(G_{n-k,k}) \oplus \xi \otimes_K \xi^* = n\xi_R$$

where $n\xi_R$ *denotes the n-fold Whitney sum* $\xi_R \oplus \cdots \oplus \xi_R$.

Let $\zeta = \xi_{k,n-k}$ be the canonical K-vector bundle over $G_{k,n-k}$. Since $G_{n-k,k} = G_{k,n-k}$, we have

$$\tau(G_{n-k,k}) \oplus \zeta \otimes_K \zeta^* = n\zeta_R.$$

Combining this equation with the equation above and noting that $\xi \oplus \zeta$ is trivial, we obtain

COROLLARY 1.2. *If* $\xi = \xi_{n-k,k}$ *and* $\zeta = \xi_{k,n-k}$ *are the canonical bundles over* $G_{n-k,k}$, *then*

$$2\tau(G_{n-k,k}) = \xi \otimes_K \zeta^* \oplus \zeta \otimes_K \xi^*.$$

Remark 1. For $k = 1$, Theorem 1.1 is due to Atiyah [1], p. 304.

Remark 2. The analogue of Theorem 1.1 for the Grassman manifold of *oriented* k-planes in R^n is true. In fact, the proof of Theorem 1.1 given in Section 2 carries over with no changes.

A real n-plane bundle η is said to admit a *generalized almost complex structure* if there is an integer $m \geqq n/2$ such that the group of $\eta \oplus \theta^{2m-n}$ can be reduced to the unitary group $U(m)$ (θ^{2m-n} is the trivial $(2m-n)$-plane bundle). We say a manifold M admits a generalized almost complex structure if its tangent bundle does.

THEOREM 1.3. *The quaternionic Grassman manifolds admit no generalized almost complex structure (except for* $G_{1,1}(H) = S^4$).

2. Proof of Theorem 1.1. Let $V_{n,k}$ denote the Stiefel manifold of orthonormal k-frames in K^n and O_n the group of orthogonal transformations in K^n. Thus O_n is either the usual orthogonal group, the unitary group, or the sympletic group. Now, if x is an orthonormal k-frame in K^n, O_k acts on the k-dimensional subspace of K^n determined by x and, under the action, an element $g \in O_k$ takes the frame x into a frame $xg \in V_{n,k}$. In fact, this action defines a principal O_k-bundle

$$\zeta = (V_{n,k}, \pi, G_{n-k,k}).$$

Define an embedding $\phi \colon V_{n,k} \to R^{nk}$ as follows. Let $x = (x_1, \cdots, x_k)$

---

[2] John Milnor has also obtained a proof of this theorem using different techniques.

be a k-frame in $V_{n,k}$ where $x_j = (x_{1j}, \cdots, x_{nj}) \in K^n$. Identifying R^{nk} with the space $\mathfrak{M}_{n,k}$ of $n \times k$ matrices over K, we set $\phi(x) = (x_{ij})$ in $\mathfrak{M}_{n,k}$. Now, O_k acts on $\mathfrak{M}_{n,k}$ by right multiplication and determines a representation $\alpha\colon O_k \to O_{nk}$.

LEMMA 2.1. *The embedding $\phi\colon V_{n,k} \to R^{nk}$ described above is equivariant relative to the representation $\alpha\colon O_k \to O_{nk}$.*

Proof. Let $x = (x_1, \cdots, x_k)$ be an element of $V_{n,k}$ and g an element of O_k. If we represent g as a matrix (g_{ij}) relative to the basis $x_1, \cdots, x_k$, then $xg = (y_1, \cdots, y_k)$ where $y_q = \sum x_j g_{jq}$. Thus $\phi(xg) = (y_{qi})$ where $y_{qi} = \sum x_{qj} g_{ji}$ which is clearly the matrix product $\phi(x)g$.

Let $\tau_F(\zeta)$ be the real vector bundle of vectors in $\tau(V_{n,k})$ tangent to fibers of ζ and ν_ϕ the normal bundle to the embedding ϕ. The group O_k acts on each of these bundles and by Theorem 1.1 of [7], we have

$$\tau(G_{n-k,k}) \oplus \tau_F(\zeta)/O_k \oplus \nu_\phi/O_k = \alpha(\zeta)_R$$

where $\alpha(\zeta)$ is the K-vector bundle associated with the α-extension of ζ (see [3], p. 477). It is not difficult to see that $\alpha\colon O_k \to O_{nk}$ is equivalent to the diagonal representation from which it follows that $\alpha(\zeta)_R = n\xi_R$. To complete the proof of Theorem 1.1, we need only identify the bundle $\tau_F(\zeta)/O_k \oplus \nu_\phi/O_k$ with $\xi \otimes_K \xi^*$. We now do this giving details only in the real case. Thus we agree that for the remainder of this section, the field will be the real numbers unless explicitly stated to the contrary.

Let $\mathcal{S}_k$ be the space of $k \times k$ symmetric matrices and define $A_{ij} \in \mathcal{S}_k$, $1 \leqq i \leqq j \leqq k$, by $A_{ij} = 1/\sqrt{2}(E_{ij} + E_{ji})$ for $i < j$ and $A_{ii} = E_{ii}$ where A_{ij} is the matrix with (p, q)-th entry $= \delta_{ip}\delta_{jq}$. Define maps $s_{ij}\colon V_{n,k} \to R^{nk} \approx \mathfrak{M}_{n,k}$ by $s_{ij}(x) = \phi(x)A_{ij}$.

LEMMA 2.2. *For each x in $V_{n,k}$, the vectors $s_{ij}(x)$, $1 \leqq i \leqq j \leqq k$, form a basis for the normal space to $\phi(V_{n,k})$ at $\phi(x)$. Thus, the maps s_{ij} determine $k(k+1)/2$ independent cross sections in the bundle ν_ϕ.*

Proof. Define hypersurfaces L_{ij}, $1 \leqq i \leqq j \leqq k$, in R^{nk} by

$$L_{ij} = \{(x_{pq}) \in \mathfrak{M}_{n,k} \ni \sum_r x_{ri} x_{rj} = \delta_{ij}\}$$

It is a matter of elementary calculus to show that $(x_{pq})A_{ij}$ is normal to L_{ij} at (x_{pq}). Furthermore, $\phi(V_{n,k}) = \cap L_{ij}$ so the $s_{ij}(x)$ describe normals to $\phi(V_{n,k})$. Since the A_{ij} form a base for $\mathcal{S}_k$, it follows that the $s_{ij}(x)$ form a basis for the normal space to $\phi(V_{n,k})$ at $\phi(x)$.

Define a representation $\beta\colon O_k \to O_{k(k+1)/2}$ by requiring that $a\beta(g) = gag^{-1}$ for $g \in O_k$, $a \in \mathcal{S}_k \approx R^{k(k+1)/2}$.

Lemma 2.3. *The bundle ν_ϕ/O_k is equivalent to $\beta(\zeta)$.*

Proof. Let $[s_{ij}(x)]$ denote the frame determined by the $s_{ij}(x)$ (in some fixed ordering). The map $s(x) = [s_{ij}(x)]$ determines a cross section in the principal bundle η associated with ν_ϕ. Let $X = \{(x,y) \in E_\eta \times E_\eta \ni \pi_\eta x = \pi_\eta y\}$ and $f\colon X \to O_{k(k+1)/2}$ the map with $x = yf(x,y)$ for $(x,y) \in X$. Then, for $g \in O_k$,

$$\begin{aligned} s(xg)g^{-1} &= [xgA_{ij}g^{-1}] \\ &= [xA_{ij}\beta(g)] \\ &= [s_{ij}(x)]\beta(g) \end{aligned}$$

so $f(s(xg)g^{-1}, s(x)) = \beta(g)$. Lemma 2.3 now follows from Lemma 6.1 of [7].

We now turn our attention to the bundle $\tau_F(\zeta)/O_k$.

Let $\mathcal{T}_k$ denote the space of $k \times k$ skew symmetric matrices with basis $B_{ij} = E_{ij} - E_{ji}$, $1 \leqq i \leqq j \leqq k$ and define maps $t_{ij}\colon V_{n,k} \to R^{nk}$ by $t_{ij}(x) = \phi(x)B_{ij}$. Identifying the fibers in $\tau_F(\zeta)$ with subspaces of R^{nk} via the embedding ϕ, we have

Lemma 2.4. *The mappings $t_{ij}\colon V_{n,k} \to R^{nk}$ determine $k(k-1)/2$ independent cross sections in the bundle $\tau_F(\zeta)$.*

Proof. Suppose, first of all, that η is an arbitrary principal G bundle and that $\mu\colon E_\eta \times G \to E_\eta$ is the principal map. Then, for each $x \in E_\eta$, we have a map $\mu_x\colon G \to E_\eta$, $\mu_x(g) = \mu(x,g)$, and therefore a map $d\mu_x$ from the tangent space G_e of G at the identity to the tangent space of E_η at x. In fact, $d\mu_x$ takes G_e isomorphically into the space of tangents along fibers at x. Thus, if $V_1, \cdots, V_r$ is a basis for G_e, the maps $x \to d\mu_x(V_i)$ determine r-independent cross sections in the bundle $\tau_F(\eta)$.

In our case, the group of the bundle is O_k and its tangent space at the identity can be identified with $\mathcal{T}_k$. If the basis B_{ij}, $1 \leqq i < j \leqq k$, is used with the method above, it is readily seen that the resulting cross sections are exactly t_{ij}.

Identifying $R^{k(k-1)/2}$ with the space $\mathcal{T}_k$, we define a representation $\gamma\colon O_k \to O_{k(k-1)/2}$ by requiring that $a\gamma(g) = gag^{-1}$ for $a \in \mathcal{T}_k$, $g \in O_k$. Just as in Lemma 2.3, this representation is exactly what is needed to prove

LEMMA 2.5. *The bundle* $\tau_F(\zeta)/O_k$ *is equivalent to* $\gamma(\zeta)$.

Combining Lemmas 2.3 and 2.5, we know that

$$\tau_F(\zeta)/O_k \oplus \nu_\phi/O_k = (\beta \oplus \gamma)(\zeta)$$

where $\beta \oplus \gamma$ denotes the direct sum of the representations β and γ. Now, $R^{k^2} \approx \mathfrak{M}_{k,k} \approx \mathcal{S}_k \oplus \mathcal{T}_k$ and the representation $\beta \oplus \gamma$ can be identified with $\lambda\colon O_k \to O_{k^2}$ where $a\lambda(g) = gag^{-1}$, $g \in O_k$, $a \in \mathfrak{M}_{k,k}$. It is easily checked that λ is equivalent to the tensor representation. This completes the proof of Theorem 1.1.

3. Proof of Theorem 1.3.

We agree that, throughout this section, the underlying field will be the field of quaternions unless explicitly stated to the contrary.

LEMMA 3.1. *The manifold* $G_{n,k}$ *admits a generalized almost complex structure if and only if the bundle* $\xi \otimes_H \xi^*$ *admits a generalized almost complex structure.*

Proof. For any finite complex X, we let $KO(X)$ denote the ring of stable real vector bundles over X and $KU(X)$ the ring of stable complex vector bundles over X (see Atiyah-Hirzebruch [2]). Let $\rho\colon KU(X) \to KO(X)$ denote the additive homomorphism defined by the map $\xi \to \xi_R$. It is easily seen that a bundle ζ has a generalized almost complex structure if and only if ζ (as an element of $KO(X)$ is in the image of ρ.

Now, let $X = G_{n,k}$. We know from Theorem 1.1 that $\tau(G_{n,k}) = (n+k)\xi_R - \xi \otimes_H \xi^*$ as an element of $KO(G_{n,k})$. Since the group of ξ_R can be reduced to $Sp(k)$, $(n+k)\xi_R$ certainly has a generalized almost complex structure. However, ρ is additive so $(n+k)\xi_R - \xi \otimes_H \xi^*$ is in the image of ρ if and only if $\xi \otimes_H \xi^*$ is in the image of ρ. This proves Lemma 3.1.

The proof of Theorem 1.3 proceeds by induction. For $n = 1$, $G_{n,k}$ is a quaternionic projective space which is known to admit no generalized almost complex structure for $k > 1$ (see Hirzebruch [4] and Massey [6]). Suppose that, for all k, $G_{n-1,k}$ admits no generalized almost complex structure. Then, by Lemma 3.1, $\xi_{n-1,k} \otimes \xi^*_{n-1,k}$ admits no generalized almost complex structure. Now, if $j\colon G_{n-1,k} \to G_{n,k}$ is the inclusion map, $\xi_{n-1,k} = j^*\xi_{n,k}$ (the bundle induced by j from $\xi_{n,k}$) so $\xi_{n-1,k} \otimes_H \xi^*_{n-1,k} = j^*(\xi_{n,k} \otimes_H \xi^*_{n,k})$. Thus, $\xi_{n,k} \otimes_H \xi^*_{n,k}$ can admit no generalized almost complex structure and the theorem is proved.

4. Appendix: The Pontrjagin class of a quaternionic tensor product.

In this section, we derive an expression for the Pontrjagin class of the quaternionic tensor product of two bundles. We first recall the splitting principal (see Hirzebruch [5], p. 97).

Let ζ be a quaternionic n-plane bundle. Associated with ζ, there is a space $X = X(\zeta)$ and a map $f\colon X \to B_\zeta$ such that

(4.1) the induced bundle $f^*\zeta$ splits into a Whitney sum of quaternionic line bundles,

(4.2) the map $f^*\colon H^*(B_\zeta\colon Z) \to H^*(X; Z)$ is a monomorphism.

We will identify $H^*(B_\zeta\colon Z)$ with its image under f^*. Thus, there are elements $x_1, \cdots, x_n \in H^4(X; Z)$ such that $p(\zeta) = \prod_{i=1}^{n} (1 + x_i)$, where $p(\zeta)$ is the total Pontrjagin class of ζ.

Suppose now that ξ and ζ are two quaternionic vector bundles over a space B with

$$p(\xi) = \prod_{i=1}^{n} (1 + x_i)$$

$$p(\zeta) = \prod_{j=1}^{m} (1 + y_j).$$

Proposition. *The total Pontrjagin class of the tensor product* $\xi \otimes_H \zeta^*$ *is given by*

$$p(\xi \otimes_H \zeta^*) = \prod_{i,j} [1 + x_i + y_j + (x_i - y_j)^2]$$

This proposition is an immediate consequence of the splitting principal and the following lemma.

Lemma 4.1. *Let ξ and ζ be quaternionic line bundles over a space B. Then*

$$p_1(\xi \otimes_H \zeta^*) = p_1(\xi) + p_1(\zeta),$$
$$p_2(\xi \otimes_H \zeta^*) = [p_1(\xi) - p_1(\zeta)]^2.$$

Proof. Define a representation $\alpha\colon Sp(1) \times Sp(1) \to SO(4)$ by requiring that $a \cdot \alpha(x, y) = xay^*$ for $x, y \in Sp(1)$ and $a \in H \approx R^4$. Then α induces a map $\lambda(\alpha)\colon B_{Sp(1)} \times B_{Sp(1)} \to B_{SO(4)}$ and, if $f, g\colon B \to Sp(1)$ are the classifying maps for ξ, ζ respectively, $\xi \otimes_H \zeta^*$ has classifying map.

$$\lambda(\alpha)(f \times g)\colon B \to B_{SO(4)}$$

Let $\gamma_1, \gamma_2\colon S^1 \to S^1 \times S^1 = T^2$ be the homomorphism defined by

$$\gamma_1(x) = (x, x), \qquad \gamma_2(x) = (x, -x).$$

Then, we have a commutative diagram

$$\begin{tikzcd} B_{T_2} \arrow[r, "\lambda(\gamma)"] \arrow[d, "i"'] & B_{T_2} \arrow[d, "j"] \\ B_{Sp(1)} \times B_{Sp(1)} \arrow[r, "\lambda(\alpha)"] & B_{SO(4)} \end{tikzcd}$$

where i and j are induced by the usual inclusions, $\lambda(\gamma)$ is induced by $\gamma = \gamma_1 \times \gamma_2$, and $\lambda(\alpha)$ is induced by α. Now, if u_1, u_2 are the canonical generators of $H^2(B_{T_2};Z)$, y_1, y_2 the canonical generators of

$$H^4(B_{Sp(1)} \times B_{Sp(1)};Z),$$

and p_1, p_2 the universal Pontrjagin classes in $H^*(B_{SO(4)};Z)$ it is readily seen that

$$\begin{aligned} \lambda(\gamma)^*u_1 &= u_1 + u_2, & \lambda(\gamma)^*u_2 &= u_1 - u_2, \\ j^*p_1 &= u_1^2 + u_2^2 & j^*p_2 &= u_1^2u_2^2, \\ i^*y_1 &= u_1^2, & i^*y_2 &= u_2^2 \end{aligned}$$

Lemma 4.1 now follows easily since i^* and j^* are monomorphisms and $f^*y_1 = p_1(\xi)$, $g^*y_2 = p_1(\zeta)$.

YALE UNIVERSITY,
NEW HAVEN, CONNECTICUT.

REFERENCES.

[1] M. F. Atiyah, "Thom complexes," *Proceedings of the London Mathematical Society* (3), vol. 11 (1961), pp. 291-310.

[2] M. F. Atiyah and F. Hirzebruch, "Vector bundles and homogeneous spaces," *Proceedings, Symposium in Pure Mathematics* (3), *Differential Geometry* (1961), pp. 7-38.

[3] A. Borel and F. Hirzebruch, "Characteristic classes and homogeneous spaces I," *American Journal of Mathematics*, 80 (1958), pp. 458-538.

[4] F. Hirzebruch, "Über die Quaternionalen Projektiven Räume," *Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Mathematik-Naturwissenschaftliche Klasse*, 1954, pp. 301-312 (1954).

[5] ———, "Neue Topologische Methoden in der Algebraischen Geometrie," *Ergebnisse der Mathematik und Grenzgebiete* (N. F.), 9, Berlin, 1956).

[6] W. S. Massey, "Non-existence of almost-complex structures on quaternionic projective spaces," *Pacific Journal of Mathematics*, vol. 12 (1962), pp. 1379-1384.

[7] R. H. Szczarba, "On tangent bundles of fiber spaces and quotient spaces," *American Journal of Mathematics*, vol. 86 (1964), pp. 685-697.

GEODESIC PARALLEL COORDINATES IN THE LARGE.

By Philip Hartman.[1]

In the first part of [4], Fiala deals with the following: Consider a complete, analytic, Riemann element of arclength defined on a plane M; $\mathcal{B}$ a closed, analytic, Jordan curve on M; and $\gamma(r)$ the set of points z inside of $\mathcal{B}$ [or outside of $\mathcal{B}$] at a distance r from $\mathcal{B}$, $\operatorname{dist}(\mathcal{B},z)=r$. (Here $\operatorname{dist}(\mathcal{B},z)$ is defined in terms of arcs of least length joining $\mathcal{B}$ and z.) Fiala shows that there exists a (possibly empty) set E of exceptional r-values which does not have a finite cluster point with the properties that (i) if $r\notin E$, then $\gamma(r)$ is the union of a finite number of Jordan curves; (ii) the length $L(r)$ of $\gamma(r)$, $r\notin E$, has a continuous derivative for which certain inequalities hold; (iii) furthermore, the function $L(r)$ can be defined at the points of $r\in E$ so as to become a continuous function. The fact (iii) greatly enhances the usefulness of Fiala's differential inequalities.

Recently, Ahlfors [1] applied Fiala's results and raised the question of their validity without assumptions of analyticity. It will be shown below that, under rather mild differentiability conditions, the analogues of (i), (ii) remain valid in the following sense: the exceptional set E is a closed set of Lebesgue measure zero (and Fiala's differential inequalities for $L(r)$ remain valid for $r\notin E$). Difficulties arise, however, with (iii); in fact, $L(r)$ can be discontinuous.

This can be seen from the example illustrated in Figure 1: Let M be a plane with the Euclidean metric, $\mathcal{B}$ the Jordan curve $ABCD\cdots$ in which ABC is a semi-circle of radius d, CD is a line segment, DE and EF are quarters of circles of radius d, FG is a line segment, etc. The heavy lines represent the "cut-locus," i.e., the endpoints of geodesics (line segments) orthogonal to $\mathcal{B}$ beyond which these geodesics do not minimize the distance from $\mathcal{B}$, the light lines indicate some of the minimizing geodesics joining $\mathcal{B}$ and the "cut-locus." (The "cut-locus" consists of a line segment and two parabolic arcs.) The closed curve just inside $\mathcal{B}$ is a locus $\gamma(r)$, $0<r<d$. The innermost closed curve is a locus $\gamma(r)$, $d<r<3d$. The locus $\gamma(d)$ consists of the dashed, closed curve and the line segments ab. It is clear

---

Received February 10, 1964.

[1] This research was supported by the Air Force Office of Scientific Research.

that the length $L(r)$ of $\gamma(r)$ has a jump discontinuity at $r = d$. (In this example, $\mathcal{C}$ is only of class C^1 but similar examples exist with $\mathcal{C}$ of class C^∞.)

It turns out that this example involves the only obstacle to the continuity of $L(r)$: the existence of "many" geodesic arcs orthogonal to $\mathcal{C}$ at both endpoints and having midpoints which are on the cut-locus but which are not "focal" points (i. e., intersections of infinitely near geodesics orthogonal to $\mathcal{C}$; Theorem 6.2 below).

In general, $L(r)$ can be extended (i. e., defined on the exceptional set E) so as to be a function which is locally of bounded variation. This function has no purely singular component. If one considers only points inside $\mathcal{C}$ [or outside $\mathcal{C}$], its purely discontinuous component is monotone in such a way that formal integrations of the differential inequalities are valid.

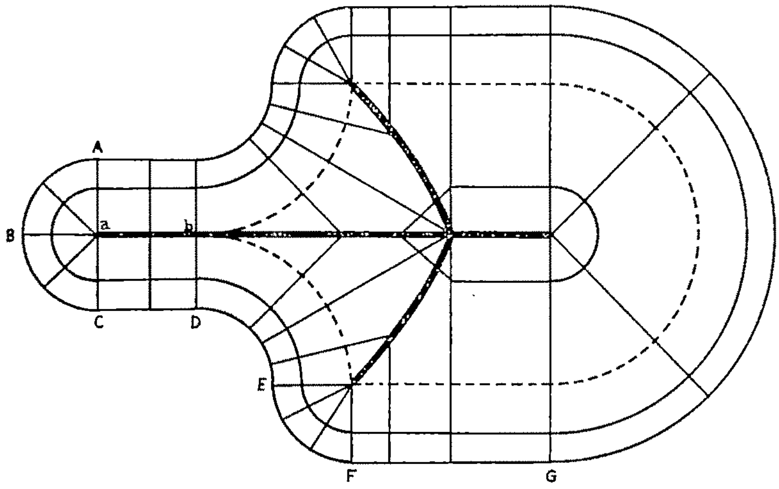

FIGURE 1.

In particular, the considerations in the second part of Fiala's work remain correct under the assumptions of this paper. Also, the arguments in Ahlfors' paper [1] are valid under the present assumptions on the Riemann metric; cf. the end of § 6 below.

This paper is concerned principally with a discussion of geodesic parallel coordinates based on a Jordan curve $\mathcal{C}$. As pointed out by Fiala ([4], p. 332), corresponding results about geodesic polar coordinates follow by choosing $\mathcal{C}$ to be a small geodesic parallel circle.

For convenience, references to known facts about the "focal points" and "extreme points" defined below will be made to Fiala [4], although many of them were given earlier by Poincaré [11], Myers [10], and Whitehead [12].

1. The Riemann manifold *M*. Let $E(x,y)$, $F(x,y)$, $G(x,y)$ be defined for all (x,y) and consider the element of arclength

$$ds^2 = Edx^2 + 2Fdxdy + Gdy^2. \tag{1.1}$$

The "local" assumptions on (1.1) are (i) ds^2 is positive definite; (ii) E, F, G are of class C^2; (iii) ds^2 has a curvature $K(x,y)$ of class C^1. The "in the large" assumption is that (iv) ds^2 is complete (in the sense of Hopf and Rinow [9]; cf. below). Local condition (iii) holds if E, F, G are of class C^3 or if (1.1) is the element of arclength on a surface of class C^3 embedded in Euclidean space. (Remarks about reducing the local assumptions (ii), (iii) are made at the end of § 2.)

Below M will refer to the Riemann manifold of points (x,y) and the element of arc-length (1.1). For brevity, write z in place of (x,y). If $\mathcal{A}: z(t) = (x(t), y(t))$ for $a \leqq t \leqq b$ is a rectifiable arc and $z(t)$ is absolutely continuous, then "length of $\mathcal{A}$" will mean the Riemann length,

$$L(\mathcal{A}) = \int_a^b \{Ex'^2 + 2Fx'y' + Gy'^2\}^{\frac{1}{2}} dt, \tag{1.2}$$

where the argument of E, F, G is $z(t)$. By the distance between two points $z_0 = (x_0, y_0)$, $z_1 = (x_1, y_1)$ is meant $d(z_0, z_1) = \inf L(\mathcal{A})$, where the "inf" refers to all rectifiable arcs $\mathcal{A}$ joining z_0 and z_1. Then M becomes a metric space with the metric $d(z_0, z_1)$. The completeness of (1.1) means that M is a complete metric space; [9].

If $\mathcal{A}$ is a rectifiable arc joining z_0, z_1 such that $L(\mathcal{A}) = d(z_0, z_1)$, then $\mathcal{A}$ will be called a *minimizing arc* (which will be abbreviated to *min-arc*).

PROPOSITION 1.1. *If z_0, z_1 are distinct points, then there exist min-arcs $\mathcal{A}$ joining them and any min-arc $\mathcal{A}$ is a geodesic*; [9].

2. Geodesic parallel coordinates. The considerations below deal with geodesic parallel coordinates based on a Jordan curve

$$\mathcal{A}: z_0(q) = (x_0(q), y_0(q)) \tag{2.1}$$

of class C^2. Let $L_0 = L(\mathcal{B})$, q is (Riemann) arclength on $\mathcal{B}$, and $z_0(q)$ is defined for $0 \leqq q \leqq L_0$. When convenient below, $z_0(q)$ and other functions of q will be considered periodic of period L_0.

For a fixed q, let $\mathcal{G}(q): z(p,q) = (x(p,q), y(p,q))$ be the unique geodesic starting (for $p = 0$) at the point $z_0(q) \in \mathcal{B}$, orthogonal to $\mathcal{B}$, $z(p,q)$ is outside of $\mathcal{B}$ for small $p > 0$, finally, p is (Riemann) arclength on $\mathcal{G}(q)$. Then, by local condition (iii) on ds^2, $z(p,q)$ is of class C^2; [6]. The completeness of (1.1) implies that $z(p,q)$ is defined for all p, $-\infty < p < \infty$; [9].

In parallel geodesic "coordinates" (p, q), (1.1) becomes

$$ds^2 = dp^2 + f^2(p, q)\,dq^2, \tag{2.2}$$

where, in terms of $x_p = \partial x/\partial p$, $x_q = \partial x/\partial q, \cdots$,

$$\begin{aligned} &Ex_p^2 + 2Fx_py_p + Gy_p^2 = 1\\ &Ex_px_q + F(x_py_q + y_px_q) + Gy_py_q = 0.\\ &Ex_q^2 + 2Fx_qy_q + Gy_q^2 = f^2(p, q). \end{aligned} \tag{2.3}$$

The argument of E, F, G in these relations is $z(p, q)$. The function $f^2(p, q)$ in (2.3) has a square root of class C^1 satisfying

$$\begin{aligned} &f(0, q) = 1,\\ &f_p(0, q) = \kappa(q) = \text{geodesic curvature of } \mathcal{B} \text{ at } z_0(q),\\ &f(p, q) = (EG - F^2)^{\frac{1}{2}}\,\partial(x, y)/\partial(p, q). \end{aligned} \tag{2.4}$$

Also, $f(p, q)$ has a continuous second partial derivative with respect to p satisfying the Jacobi equation

$$f_{pp}(p, q) + K(z(p,q))f(p, q) = 0. \tag{2.5}$$

Note that, at a point (p, q),

$$f(p, q) = 0 \Longleftrightarrow z_q(p, q) = 0 \Longleftrightarrow \partial(x, y)/\partial(p, q) = 0. \tag{2.6}$$

If $f(p_0, q_0) = 0$, then, since $f \not\equiv 0$, (2.5) implies that $f_p(p_0, q_0) \neq 0$. Hence, $f(p, q) = 0$ has a solution $p = p(q)$ of class C^1 for q near q_0, $p(q_0) = p_0$, and

$$dp(q)/dq = -f_q(p(q), q)/f_p(p(q), q). \tag{2.7}$$

A p-value where $f(p, q) = 0$, or the corresponding point $z(p, q) \in M$, will be called a *focal point* on the geodesic $\mathcal{B}(q)$.

If $f(p, q) = 0$ for a positive p-value, let $p = p(q)$ denote the least such p-value; otherwise, let $P(q) = \infty$. Correspondingly, let $p = N(q)$ be the largest negative zero of $f(p, q)$ or let $N(q) = -\infty$ if there are no such zeros. The point $p = P(q)$ [or $p = N(q)$] will be called the first positive [or negative] focal point on $\mathcal{B}(q)$.

Remarks on local conditions. If conditions (ii), (iii) are replaced by (ii′) E, F, G are of class C^1 and (iii′) ds^2 has a continuous curvature $K(x, y)$ (in the sense that the formula of Gauss-Bonnet is valid), then $\mathcal{B}(q)$ is unique, $z(p, q)$ and $z_p(p, q)$ are of class C^1, $f(p, q)$ is continuous and has continuous partial derivatives f_p, f_{pp} satisfying (2.4), (2.5); cf. [5] or [6]. Correspondingly, local solutions $p = p(q)$ of $f(p,q) = 0$ are continuous (when they exist.) Under these conditions, the results of § 6, in particular, the main Theorems 6.1 and 6.2, remain valid; however, it is not clear, in this case, that the set of exceptional r-values is of Lebesgue measure zero, i. e., that (6.3)-(6.4) hold almost everywhere. (It can be mentioned that the results of § 5 are valid with the following possible exceptions: the parts of Proposition 5.4 involving $\rho(q_0) = P(q_0) < \infty$ and $\nu(q_0) = N(q_0) > -\infty$; Corollary 5.1; the assertion "of Lebesgue measure zero" in Corollary 5.3 and Lemma 5.2; "C^2" should be replaced by "C^1" in (i) (ii) of Proposition 5.6; Lemma 5.4.)

3. **Preliminaries.** For reference, it will be convenient to state here some notations and tools from the theory of functions of a real variable to be used below in avoiding assumptions of analyticity.

If S is a Lebesgue measurable set, $|S|$ denotes its Lebesgue measure.

Let $h(q)$ be a continuous function of bounded variation on an interval $q_1 \leqq q \leqq q_2$ and $T_h(q)$ a non-decreasing function such that $T_h(\tau) - T_h(\sigma)$ is the total variation of $h(q)$ on $[\sigma, \tau] \subset [q_1, q_2]$. If S is a Borel subset of $[q_1, q_2]$, "variation of h over S" means the Lebesgue-Stieltjes measure $\int_S |dh| = \int_S dT_h(q)$ of S or, equivalently, the Lebesgue measure $|T_h(S)|$ of the image of S under the map $q \to T_h(q)$.

Lemma 3.1. (Banach [2], p. 228). *Let $h(q)$ be a continuous function on $q_1 \leqq q \leqq q_2$ and $n(r) = n_h(r)$ the number, $0 \leqq n(r) \leqq \infty$, of solutions q of the equation $h(q) = r$. Then $h(q)$ is of bounded variation if and only if $n(r)$ is Lebesgue integrable over $-\infty < r < \infty$; in which case,*

$$\int_{-\infty}^{+\infty} n(r)\,dr = \{\text{var}\, h \text{ on } [q_1, q_2]\} = |T_h([q_1, q_2])| \tag{3.1}$$

and $0 \leqq n(r) < \infty$ almost everywhere (i. e., except for a set of r-values of Lebesgue measure zero).

In §§ 5-6 below, the following consequence of this fact will be needed:

Corollary 3.1. *Let $h(q)$ be a continuous function of bounded variation on $[q_1, q_2]$ and $V(r)$ the variation of h over the set $\{q: h(q) < r\}$. Then $V(r)$ is absolutely continuous and*

$$V(r) = \int_{-\infty}^{r} n(u)\,du, \tag{3.2}$$

so that

$$V(\tau) - V(\sigma) = \int_\sigma^\tau n(u)\,du. \tag{3.3}$$

The following well known fact will also be used.

Proposition 3.1. Let $h(q)$ be continuous and of bounded variation on $[q_1, q_2]$. Then the variation of h over the set $\{q\colon dh/dq$ exists and is $0\}$ is zero.

4. Locus of first focal points. A standard theorem (Jacobi) in the calculus of variations is the following:

Proposition 4.1. When $P(q)$ [or $N(q)$] is finite-valued, the locus of first positive [or negative] focal points $z = z(P(q), q)$ [or $z = z(N(q), q)$] is an envelope of the family of geodesics $z = z(p, q)$. It is of class C^1 on any q-interval on which $P(q)$ [or $N(q)$] is finite-valued and $dP/dq \neq 0$ [or $dN/dq \neq 0$]; furthermore, the Riemann arclength element along such an arc is $ds = |dP(q)|$ (or $ds = |dN(q)|$).

In fact, $dz(P(q), q)/dq = z_p(P, q)\,dP/dq$.

LEMMA 4.1. *For any $r > 0$, the functions $P_r(q) = \min(P(q), r)$, $N_r(q) = \max(N(q), -r)$ are uniformly Lipschitz continuous.*

Proof. Consider only $P(q)$. If $P(q_0) < \infty$, then $P(q)$ is of class C^1 for q near q_0 and (2.7) holds with $p(q) = P(q)$. Note that $f_q(p, q)$ is bounded on the set $\{(p, q): |p| \leqq r, 0 \leqq q \leqq L_0\}$ and that $1/f_p(p, q)$ is also bounded on the set $\{(p, q): f(p, q) = 0, |p| \leqq r\}$. Hence there exists a constant $c_0 = c_0(r)$ such that

$$|P(q_1) - P(q_2)| \leqq c_0 |q_1 - q_2| \text{ if } 0 < P(q) \leqq r \text{ on } [q_1, q_2].$$

Since the constants c_0 depends only on r and not on the interval $[q_1, q_2]$, the assertion follows.

COROLLARY 4.1. *For almost all r, the number of q-values, $0 \leqq q \leqq L_0$, where $P(q) = r$ [or $N(q) = r$] is finite.*

This is a consequence of Lemma 4.1 and Lemma 3.1. For if $n_s(r)$ is the number of q-values where $P_s(q) = \min(P(q), s)$ takes the value r, then $n_s(r) < \infty$ for almost all r. But if $r < s$, then $n_s(r)$ is the number of q-values where $P(q) = r$.

COROLLARY 4.2. *Let Z_0 be the set of numbers r such that $P(q) = r$ or*

$N(q) = r$ has at least one solution $q = q_0$ such that dP/dq or dN/dq exists and is zero at $q = q_0$. Then Z_0 is of Lebesgue measure zero.

Cf. Proposition 3.1.

5. Extreme points. For any point $z = (x, y)$, let $|d(z)| = \inf L(\mathcal{A})$, where the "inf" refers to all rectifiable Jordan arcs joining a point of $\mathcal{C}$ and z, and let $d(z) < 0$, $= 0$, or > 0 according as z is inside, on, or outside $\mathcal{C}$. If $\mathcal{A}$ has the property that it joins a point $z_0(q) \in \mathcal{C}$ and z and $|d(z)| = L(\mathcal{A})$, then $\mathcal{A}$ will be called a $\mathcal{C}$-min-arc.

Proposition 5.1. For any point $z \notin \mathcal{C}$, there exists at least one $\mathcal{C}$-min-arc $\mathcal{A}$. If $\mathcal{A} = \widehat{z_0(q)z}$ is a $\mathcal{C}$-min-arc, then $\mathcal{A}$ is a min-arc and a sub-arc of $\mathcal{G}(q)$. If z_1 is a point of $\mathcal{A}$, then the sub-arc $\widehat{z_0(q)z_1}$ of $\mathcal{A}$ is a $\mathcal{C}$-min-arc and the sub-arc $\widehat{z_1 z}$ is a min-arc.

Propositions 1.1 and 5.1 imply the following:

Proposition 5.2. If $\mathcal{A}_1$, $\mathcal{A}_2$ are distinct [$\mathcal{C}$-min-arcs] min-arcs, then $\mathcal{A}_1$, $\mathcal{A}_2$ have at most their endpoints [not on $\mathcal{C}$] in common.

Let $\rho(q) = \sup p$ taken over positive p-values such that the sub-arc of $\mathcal{G}(q)$ corresponding to $[0, p]$ is a $\mathcal{C}$-min-arc; correspondingly, let $\nu(q) = \inf p$ taken over all $p < 0$ such that the sub-arc of $\mathcal{G}(q)$ belonging to $[p, 0]$ is a $\mathcal{C}$-min-arc. Thus $-\infty \leqq \nu(q) < 0 < \rho(q) \leqq \infty$. A finite p-value $p = \rho(q)$, $\nu(q)$ or corresponding point $z(\rho(q), q)$, $z(\nu(q), q)$ of M will be called an *extreme point* of $\mathcal{G}(q)$. Following Whitehead [12], the set of extreme points $z(\rho(q), q)$, $z(\nu(q), q)$ can be called the "cut-locus" of M relative to $\mathcal{C}$.)

Proposition 5.3. An extreme point $z \in M$ is either a first focal point on a $\mathcal{C}$-min-arc or the intersection of at least two $\mathcal{C}$-min-arcs; cf., e.g., [4], p. 314.

Proposition 4.1 (i.e., the "Envelope Theorem") implies

Proposition 5.4. $\rho(q) \leqq P(q)$ for all q, while $\rho(q_0) = P(q_0) < \infty$ implies that $dP(q_0)/dq$ exists and is 0; similarly, $\nu(q) \geqq N(q)$ for all q, while $\nu(q_0) = N(q_0) > -\infty$ implies that $dN(q_0)/dq$ exists and is 0.

For the sake of brevity, only the set of extreme points outside of $\mathcal{C}$ and $\rho(q)$ will be considered in the remainder of this section. It will be clear that analogous results hold for extreme points inside $\mathcal{C}$ and $\nu(q)$.

COROLLARY 5.1. *If* $\rho(q_0)=P(q_0)<\infty$, *then the lower left and upper right derivatives of* $\rho(q)$ *satisfy* $D_L\rho(q_0)\geqq 0\geqq D^R\rho(q_0)$,

For if q_1 is near q_0, then $\rho(q_1)-\rho(q_0)\leqq P(q_1)-P(q_0)$ by Proposition 4.1.

Proposition 5.5. The function $\rho(q)$ is continuous (even at points where it is not finite-valued).

Cf. Whitehead [12], pp. 701-702. (Fiala's similar proof [4], pp. 316-317, makes unnecessary use of his analyticity assumptions.)

Let a point z_0 outside of $\mathcal{C}$ be called ***normal*** if it is the endpoint of exactly two $\mathcal{C}$-min-arcs and is not a first focal point on either of them. In this case, $z_0=z(\rho(q),q)$ holds for exactly two values of q, say $q=q_1,q_2$ subject to $0\leqq q_1<q_2<L_0$. Furthermore, $\rho(q_j)<P(q_j)$ for $j=1,2$. A point $z(\rho(q),q)$ which is not normal will be called *anormal.*

LEMMA 5.1. *Let* $z_1,z_2,\cdots$ *be a sequence of distinct anormal points and let* $z_n\rightarrow z_0$ *as* $n\rightarrow\infty$. *Then* z_0 *is a first focal point on a* $\mathcal{C}$*-min-arc.*

Proof. If infinitely many of the points $z_1,z_2,\cdots$ are first focal points of $\mathcal{C}$-min-arcs, then it is clear that z_0 is also.

Thus, it can be supposed that no z_n is a first focal point of a $\mathcal{C}$-min-arc. Hence, z_n is the endpoint of at least three $\mathcal{C}$-min-arcs. Thus, for fixed n, there exist three q-values $q_{1n}<q_{2n}<q_{3n}$ $(<q_1+L_0)$ such that $z(\rho(q_{in}),q_{in})=z_n$ for $i=1,2,3$. It will be supposed that the enumeration of q_{1n},q_{2n},q_{3n} is chosen so that $0\leqq q_{1n}<L_0$ and that the $\mathcal{C}$-min-arcs on $\mathcal{G}(q_{1n})$, $\mathcal{G}(q_{3n})$ together with the arc on $\mathcal{C}$ corresponding to $q_{1n}\leqq q\leqq q_{3n}$ forms a Jordan curve with an interior I_n containing the $\mathcal{C}$-min-arc on $\mathcal{G}(q_2)$. The sets I_n and I_m are either disjoint or one contains the other.

It can be supposed that $q^i=\lim q_{in}$ exists, as $n\rightarrow\infty$, for $i=1,2,3$. If two of these limits are equal $(\operatorname{mod} L_0)$, say $q^1=q^2$ [or $q^2=q^3$ or $q^3=q^1+L_0$], then it is clear that the corresponding limit point is a first focal point on the $\mathcal{C}$-min-arc on $\mathcal{G}(q^1)=\mathcal{G}(q^2)$ [or $\mathcal{G}(q^2)=\mathcal{G}(q^3)$ or $\mathcal{G}(q^3)=\mathcal{G}(q^1)$].

Thus, in order to complete the proof it is sufficient to show that $q^1<q^2<q^3<q^1+L_0$ is impossible.

If infinitely many of the sets $I_1,I_2,\cdots$ are disjoint, then it is clear that $q^1=q^2=q^3$. Thus it can be supposed that $I_1\supset I_2\supset\cdots$ or that $I_1\subset I_2\subset\cdots$. Assume the first case. Let I_{nR}, I_{nL} denote the portion of I_n contained, respectively, inside the Jordan curve formed by the $\mathcal{C}$-min-arcs on $\mathcal{G}(q_1)$, $\mathcal{G}(q_2)$ and an arc of $\mathcal{C}$ or by the $\mathcal{C}$-min-arcs on $\mathcal{G}(q_2)$, $\mathcal{G}(q_3)$

and an arc of $\mathcal{B}$. Either $I_{n+1} \subset I_{nL}$ or $I_{n+1} \subset I_{nR}$ holds for infinitely many values of n. For the sake of definiteness, let $I_{n+1} \subset I_{nR}$ hold for infinitely many n; i.e., $q_{1n} < q_{1,n+1} < q_{2,n+1} < q_{3,n+1} < q_{2n} < q_{3n}$ holds for infinitely many n. Then $q^2 = q^3$. This completes the proof, since the case $I_1 \subset I_2 \subset \cdots$ is similar.

COROLLARY 5.2. *An anormal point z_0 is either a first focal point on a $\mathcal{B}$-min-arc or an isolated anormal point. In particular, the set of anormal points z_0 is closed.*

A number $r > 0$ will be called *anormal* if there exists a value of $q = q_0$ such that $\rho(q_0) = r$ and $z(r, q_0)$ is anormal.

COROLLARY 5.3. *The set of anormal r-values is closed and of Lebesgue measure zero.*

Proof. In fact, it is clear from the last corollary and from Proposition 5.4 that an anormal r-value is either contained in the set Z_0 of Corollary 4.2 or is an isolated anormal value. Hence the set of anormal values is of Lebesgue measure zero. Lemma 5.1 implies that this set is closed.

Proposition 5.6. Let z_0 be a normal point (so that $z_0 = z(\rho(q), q)$ has exactly two solutions $q = q_1, q_2 \pmod{L_0}$, $q_1 < q_2 < q_1 + L_0$ and $\rho(q_j) < P(q_j)$ for $j = 1, 2$). Then (i) $\rho(q)$ is of class C^2 on a neighborhood of $q = q_1, q_2$; (ii) there exists a strictly monotone function $v(q)$ of class C^2 on a neighborhood of $q = q_2$ with the properties that $v(q_2) = q_1$, $\rho(v(q)) = \rho(q)$ and $z(\rho(v(q)), v(q)) = z(\rho(q), q)$ for q near q_2; finally, (iii) the locus of extreme points $z = z(\rho(q), q)$ for q near q_2 (or q_1) bisects the angle between the $\mathcal{B}$-min-arcs on $\mathfrak{B}(q)$, $\mathfrak{B}(v(q))$ meeting at $z(\rho(q), q)$.

Cf. [4], pp. 317-319. The proof follows from the fact that $z = z(p, q)$ has two inverses of class C^2 near z: $p = p_j(z)$, $q = q_j(z)$ such that $p_j(z_0) = \rho(q_1) = \rho(q_2)$ and $q_j(z_0) = q_j$ for $j = 1, 2$. The normality of z_0 and the fact that the set of anormal points is closed imply that the locus of extreme points is determined by the equation $p_1(z) = p_2(z)$. Since $\mathfrak{B}(q_1)$, $\mathfrak{B}(q_2)$ cannot meet at a zero angle, the implicit function theorem is applicable (i.e., either $\partial(p_1 - p_2)/\partial x$ or $\partial(p_1 - p_2)/\partial y$ does not vanish at z_0).

Proposition 5.7. Let z_0, q_1, q_2 be as in Proposition 5.6. Then $d\rho/dq = 0$ at $q = q_1$ (and/or $q = q_2$) if and only if $\mathfrak{B}(q_1)$, $\mathfrak{B}(q_2)$ are continuations of each other (i.e., the two $\mathcal{B}$-min-arcs meeting at z_0 meet at an angle π).

Since (p, q) is an orthogonal coordinate system on M, this proposition follows from (iii) in Proposition 5.6 and the fact that the arc $p=\rho(q)$ is orthogonal to the line $q=q_1$ if and only if their images on M are orthogonal at z_0.

(In non-trivial analytic cases, the possibility that the $\mathcal{C}$-min-arcs on $\mathcal{G}(q_1)$, $\mathcal{G}(q_2)$ are continuations of each other can occur only for isolated $\mathcal{C}$-min-arcs; [4], pp. 320-322.)

Proposition 5.8. Let z_0 be an isolated anormal point which is not a first focal point on any $\mathcal{C}$-min-arc through it. Then (i) there exists only a finite number n $(\geqq 3)$ of q-values $(\operatorname{mod} L_0)$, say, $q=q_1,\cdots,q_n$, such that $z_0=z(\rho(q),q)$; (ii) $\rho(q)$ is of class C^1 on all small closed intervals ending at q_j (but $d\rho(q+0)/dq\neq d\rho(q-0)/dq$ at $q=q_j$ is not excluded); finally (iii) the set of extreme points in a neighborhood of z_0 consists of n arcs of class C^1 ending at z_0; at each point z of this set, these arcs bisect the angle between a pair of $\mathcal{C}$-min-arcs intersecting at z.

Proof. Assertion (i) is clear since z_0 is not a first focal point on any $\mathcal{C}$-min-arc ending at z_0. For any q_j and small q-interval I_1 abutting at q_j, there is a q_i, $i=i(j,I_1)$, and an interval I_2 abutting at q_i such that there exists a continuous, one-to-one map $v\colon I_2\rightarrow I_1$, and $\rho(v(q))=q$, $z(\rho(v(q)),v(q))=z(\rho(q),q)$ for q on I_2. Also, the locus $z=z(\rho(q),q)$ for $q_i\neq q\in I_2$ bisects the angle between $\mathcal{G}(q)$, $\mathcal{G}(v(q))$. The assertions (ii), (iii) follow from a continuity argument. (The finiteness of the one sided derivatives of $\rho(q)$ at q_j follows from the fact that the angle between $\mathcal{G}(q)$, $\mathcal{G}(v(q))$ cannot be zero.).

A number $r>0$ will be called *exceptional* if it is either anormal or if it is normal but there exists a q-value q_0 such that $\rho(q_0)=r$, $d\rho(q_0)/dr=0$.

LEMMA 5.2. *The set of exceptional r-values is closed and of Lebesgue measure zero.*

Proof. The set is closed since the set of anormal points is closed and $d\rho/dq$ is continuous at normal points.

In view of Corollaries 4.2 and 5.2, it is sufficient to show that the set of normal exceptional r-values is of Lebesgue measure 0. The set of q-values on $0<q<L_0$ for which $z(\rho(q),q)$ is normal is open, by Corollary 5.2. Let $[\alpha,\beta]$ be a closed q-interval such that $q\in[\alpha,\beta]$ implies that $z(\rho(q),q)$ is normal. Then $d\rho/dq$ is continuous on $[\alpha,\beta]$. Hence the set of r-values such that $\rho(q)=r$ has a solution $q\in[\alpha,\beta]$ at which $d\rho(q)/dq=0$ is of Lebesgue measure zero; cf. Proposition 3.1. Thus Lemma 5.2 follows.

LEMMA 5.3. *Let $r > 0$ be non-exceptional. Then the equation $\rho(q) = r$ has at most a finite number $n_\rho(r)$ of solutions q $(\operatorname{mod} L_0)$. The integer $n_\rho(r)$ is even and constant on any interval of non-exceptional r-values.*

Proof. If $n_\rho(r) = \infty$, then there exists a q-value q_0 such that $\rho(q_0) = r$ and q is a limit point of q-values satisfying $\rho(q) = r$. But this contradicts the fact that $z(\rho(q_0), q_0)$ is normal, $d\rho/dq$ is continuous at $q = q_0$ and $d\rho(q_0)/dq \neq 0$.

It will remain undecided whether or not the function $\rho_R(q) = \min(\rho(q), R)$ is of bounded variation. But when it is, the following holds.

LEMMA 5.4. *Let $R > 0$ be fixed, and let $\rho_R(q) = \min(\rho(q), R)$ be of bounded variation on $0 \leqq q \leqq L_0$. Then $\rho_R(q)$ is absolutely continuous.*

Proof. It can be supposed that R is non-exceptional, for otherwise consider $\rho_s(q)$ and let $s \to R - 0$.

Let $\epsilon > 0$ be fixed. Then, by Corollary 3.1, there exists an $\eta = \eta(\epsilon, R)$ such that

$$\int_\Sigma n_\rho(r)\,dr < \epsilon \text{ if } |\Sigma| < \eta \tag{5.1}$$

and Σ is a Borel subset of $[0, R]$. The integral in (5.1) is the variation of $\rho_R(q)$ over the set

$$S_1 = \{q : \rho(q) \in \Sigma\}. \tag{5.2}$$

The set of exceptional points r contained in $[0, R]$ is compact and of Lebesgue measure zero. Hence there exists a finite number of open non-overlapping intervals $\delta_1, \delta_2, \cdots, \delta_k$ containing these points such that the endpoints of $\delta_1, \cdots, \delta_k$ are non-exceptional and

$$\Sigma = \{\delta_1 \cup \cdots \cup \delta_k\} \cap [0, R] \text{ satisfies } |\Sigma| < \eta. \tag{5.3}$$

If Σ is defined by (5.3), the intervals which constitute Σ have non-exceptional endpoints. The coresponding set S_1 in (5.2) is the union of non-overlapping sub-intervals $\Delta^1, \cdots, \Delta^e$ of $[0, L_0]$ which are open relative to $[0, L_0]$. Furthermore,

$$\{\operatorname{var} \rho_R \text{ over } S_1 = \bigcup_{i=1}^{e} \Delta^i\} < \epsilon. \tag{5.4}$$

The set

$$S_2 = \{q : 0 \leqq q \leqq L_0, \rho(q) \leqq R, \rho(q) \notin \Sigma\}$$

is closed. Near each point of S_2, $\rho(q)$ is of class C^2. Hence there exists a constant $c > 0$ such that

$$(5.5) \qquad q \in S_2 \Rightarrow \mid d\rho(q)/dq \mid \leqq c.$$

Let $\Delta_1, \cdots, \Delta_j$ be an arbitrary set of non-overlapping closed intervals on $0 \leqq q \leqq L$ such that

$$(5.6) \qquad \sum_{i=1}^{j} \mid \Delta_i \mid < \epsilon/c.$$

It will be shown that if $\Delta_i = [\alpha, \beta]$ and $\Delta_i\rho_R = \rho_R(\beta) - \rho_R(\alpha)$, then

$$(5.7) \qquad \sum_{i=1}^{j} \mid \Delta_i\rho_R \mid < 2\epsilon.$$

To this end, note that the set

$$\bigcup_{i=1}^{j} \Delta_i - \bigcup_{i=1}^{e} \Delta^i$$

consists of a finite number of closed intervals, say $\Delta^{e+1}, \cdots, \Delta^m$. It is clear that if $q \in \Delta^i$ for some $i = e+1, \cdots, m$, then either $q \in S_2$ or $\rho_R(q) = R$. Let $e+1 \leqq i \leqq m$, $\Delta^i = [\alpha, \beta]$ and $\Delta^i\rho_R = \rho_R(\beta) - \rho_R(\alpha)$. If $\rho_R(\beta) = \rho_R(\alpha) = R$, then $\Delta^i\rho_R = 0$. If $\rho_R(\alpha) < R$ and $\rho_R(\beta) = R$, then $q = \alpha \in S_2$. Let $\gamma > \alpha$ be the least q-value such that $\rho(q) = R$, so that $\alpha < \gamma \leqq \beta$. Then $[\alpha, \gamma] \subset S_2$ and (5.5) imply that

$$\mid \Delta^i\rho_R \mid = \mid \rho(\gamma) - \rho(\alpha) \mid \leqq c(\gamma - \alpha).$$

This argument shows that, in all four cases, $\rho_R(\alpha) \leqq R$ and $\rho_R(\beta) \leqq R$, one has $\mid \Delta^i\rho_R \mid \leqq c \mid \Delta^i \mid$ for $e+1 \leqq i \leqq m$. Consequently,

$$\sum_{i=1}^{j} \mid \Delta_i\rho_R \mid < \int_\Sigma n_\rho(r)\,dr + c\sum_{i=e+1}^{m} \mid \Delta^i \mid < \epsilon + c\sum_{i=1}^{j} \mid \Delta_i \mid.$$

Thus (5.7) follows from (5.6). This completes the proof of the absolute continuity of $\rho_R(q)$.

6. The r-loci. Let $\bar{r} = \inf \nu(q)$, so that $-\infty < \bar{r} < 0$. Below only r-values on the range $\bar{r} \leqq r < \infty$ will be considered.

Gauss's geodesic parallels $\Gamma(r)$ are the loci $\Gamma(r) : z = z(r, q)$, $0 \leqq q \leqq L_0$. $\Gamma(r)$ is a (not necessarily simple) closed curve and $z_q(r, q_0) \neq 0$ if $r \neq P(q_0)$. Hence, for almost all r, $\Gamma(r)$ is locally an arc of class C^2.

Define $\gamma(r)$, the r-locus, to be the set of points $z \in M$ such that $d(z) = r$. Then $\gamma(r) \subset \Gamma(r)$. If $r > \bar{r}$ and $\rho(q) = r$ has no solutions, then $\gamma(r) = \Gamma(r)$ is a simple closed curve.

Proposition 6.1. If $r > 0$ is non-exceptional, then $\rho(q) = r$ has an even number of solutions $2m \pmod{L_0}$ which if $m > 0$, can be enumerated as $\alpha_1 < \beta_1 < \alpha_2 < \cdots < \beta_m$ $(< \alpha_1 + L_0)$ such that $\rho(q) > r$ on (α_k, β_k) and $\rho(q) < r$ on (β_k, α_{k+1}), for $k = 1, \cdots, m$ with $\alpha_{m+1} = \alpha_1 + L_0$. The locus $\gamma(r)$ consists of the set of points $z = z(r, q)$, $\alpha_k \leqq q \leqq \beta_k$ for $k = 1, \cdots, m$, and is a set of simple closed curves of class C^2 except for possible corners at $z(r, \alpha_k)$, $z(r, \beta_k)$; *cf.* [4], pp. 325-326. (The z-image $z = z(r, q)$ of one of the q-intervals $[\alpha_k, \beta_k]$ is, in general, only an arc on one of these closed curves.) The total (Riemann) length of these curves is

$$L(r) = \sum_{k=1}^{m} \int_{\alpha_k}^{\beta_k} f(r, q)\, dq; \tag{6.1}$$

$L(r)$ is of class C^1 and

$$dL/dr = \sum_{k=1}^{m} \{ \int_{\alpha_k}^{\beta_k} f_p(r, q)\, dq + f(r, \beta_k)\, d\beta_k/dr - f(r, \alpha_k)\, d\alpha_k/dr \}. \tag{6.2}$$

Similar results hold for non-exceptional $r < 0$.

The formulae (6.1), (6.2) are given by Fiala [4], pp. 326-328. His considerations (depending on Proposition 5.5 and the fact that $f_p(r, q)/f(r, q)$ is the geodesic curvature of the geodesic parallel $z = z(r, q)$) lead to the following differential inequalities:

THEOREM 6.1. *Let* $r \neq 0$ *be non-exceptional. Then*

$$dL(r)/dr \leqq \int_0^{L_0} \kappa(q)\, dq - C(r) \text{ if } 0 < r < \infty, \tag{6.3}$$

$$dL(r)/dr \geqq \int_0^{L_0} \kappa(q)\, dq + C(r) \text{ if } \bar{r} < r < 0, \tag{6.4}$$

where $\kappa(q)$ *is the geodesic curvature of* $\mathcal{B}$ *at* $z_0(q)$,

$$C(r) = \int\int_{\Delta(r)} K\, dA, \tag{6.5}$$

K *is the curvature of* (1.1), dA *is the Riemann element of area, and the integral in* (6.5) *is over the domain* $\Delta(r)$ *bounded by* $\mathcal{B}$ *and the r-locus* $\gamma(r)$.

Consider the (p, q)-set

$$D = \{(p, q) : \nu(q) < p < \rho(q), 0 \leqq q < L_0\}. \tag{6.6}$$

The map $(p, q) \to z(p, q)$ from D to M is one-to-one and from the closure of D to M is onto (not one-to-one). Let $\chi(p, q)$ be the characteristic function of D, so that $\chi(p,q)$ is 1 or 0 according as $\nu(q) < p < \rho(q)$ does or does not hold. Then the function (6.5) is given by

$$(6.7) \qquad C(r) = \pm \int_0^r \int_0^{L_0} K(z(p,q)) f(p,q) \chi(p,q)\, dp\, dq \text{ if } \pm r > 0,$$

The function $L(r)$ has been defined in (6.1) only for non-exceptional $r > \bar{r}$. In order to make the differential inequalities (6.3), (6.4) more useful, it will be shown that the integrated forms of the differential inequalities (6.3), (6.4) are valid, even though it may not be possible to extend $L(r)$ to a continuous function; cf. (6.11)-(6.12).

For $\bar{r} \leqq r < \infty$, put

$$(6.8) \qquad L(r) = \int_0^{L_0} \chi(r,q) f(r,q)\, dq,$$

which reduces to (6.1) for non-exceptional r.

For $\bar{r} \leqq r < \infty$, define the q-sets $Q(r)$ on $0 \leqq q \leqq L_0$ by

$$(6.9) \qquad \begin{aligned} Q(r) &= \{q : \rho(q) = r \text{ is normal, } d\rho(q)/dq = 0\},\ r > 0, \\ Q(r) &= \{q : \nu(q) = r \text{ is normal, } d\nu(q)/dq = 0\},\ r < 0. \end{aligned}$$

For $r > 0$, $q \in Q(r)$ if and only if $\rho(q) = r$, $z(r,q)$ is normal, and the two $\mathfrak{B}$-min-arcs meeting at $z(r,q)$, are arcs on the same geodesic. The sets $Q(r)$ are pairwise disjoint for positive [or negative] r-values and so, $|Q(r)| = 0$ except for an at most countable set of r-values.

For $\bar{r} \leqq r < \infty$, define the function

$$(6.10) \qquad \begin{aligned} J(r) &= \sum_{0 \leqq t \leqq r} \int_{Q(t)} f(t,q)\, dq \text{ for } r \geqq 0. \\ J(r) &= \sum_{r \geqq t \geqq 0} \int_{Q(t)} f(t,q)\, dq \text{ for } r \leqq 0, \end{aligned}$$

Since $f \geqq 0$ on the set of integration, $J(r)$ is non-decreasing for $r \geqq 0$ and non-decreasing for $r \leqq 0$. The function $J(r)$ is purely discontinuous and its discontinuity points are the r-values where $|Q(r)| \neq 0$ (for $q \in Q(r)$ implies that $f(r,q) > 0$ since $z(r,q)$ is not a first focal point on $\mathfrak{B}(q)$).

THEOREM 6.2. *The function* $H(r) = L(r) + J(r)$ *is absolutely continuous for* $\bar{r} \leqq r < \infty$.

Remark. In particular, $L(r)$ is of bounded variation on bounded intervals of $[\bar{r}, \infty)$ and has no purely singular component (in the sense of Lebesgue). Furthermore, $L(r)$ is absolutely continuous if and only if $J(r) \equiv 0$, i.e., if $|Q(r)) = 0$ for $r \geqq \bar{r}$.

COROLLARY 6.1. *The integrated forms of* (6.3), (6.4) *are valid for*

$$
\begin{aligned}
&L(b)-L(a)\leqq\int_a^b (dL/dr)\,dr \text{ for } 0\leqq a<b<\infty,\\
(6.11)\qquad &\\
&L(b)-L(a)\geqq\int_a^b (dL/dr)\,dr \text{ for } \bar{r}\leqq a<b\leqq 0;
\end{aligned}
$$

more generally, if $\phi(L)$ *is a continuously differentiable, non-decreasing function for* $0\leqq L<\infty$, *then*

$$
\begin{aligned}
&\phi(L(b))-\phi(L(a))\leqq\int_a^b \phi'(L(r))L'(r)\,dr \text{ for } 0\leqq a<b<\infty,\\
(6.12)\qquad &\\
&\phi(L(b))-\phi(L(a))\geqq\int_a^b \phi'(L(r))L'(r)\,dr \text{ for } \bar{r}\leqq a<b\leqq 0,
\end{aligned}
$$

where $\phi'(L)=d\phi/dL$ *and* $L'(r)=dL/dr$.

Proof of Theorem 6.2. Consider the case $0\leqq a<b<\infty$. Let $\epsilon>0$ be fixed and let

$$Q^\epsilon=\{q\colon \rho(q)\leqq b, f(\rho(q),q)\leqq\epsilon/2\}.$$

Then Q^ϵ is a closed subset of $0\leqq q\leqq L_0$ and can be covered by a set S^ϵ consisting of a finite number of closed subintervals of $[0,L_0]$ on which $\rho(q)<\infty$ and $f(\rho(q),q)\leqq 3\epsilon/4$. If $\rho(q)\leqq b$ and $q\notin S^\epsilon$, then $f(\rho(q),q)>\epsilon/2>0$, so that $\rho(q)<P(q)\leqq\infty$. In particular, $z(\rho(q),q)$ is either a normal or an isolated anormal point. Thus, on the closure R^ϵ of $[0,L_0]-S^\epsilon$, the function $\rho_b(q)=\min(\rho(q),b)$ has bounded right and left derivatives. The set R^ϵ consists of a finite number of closed intervals and $\rho_b(q)$ is of bounded variation on R^ϵ. Thus, by Corollary 3.1, if $0\leqq\sigma<\tau\leqq b$, the total variation of $\rho_b(q)$ over the set $\{q\colon q\in R^\epsilon, \sigma<\rho(q)\leqq\tau\}$ is given by

$$T(\{q\in R^\epsilon,\sigma<\rho(q)\leqq\tau\})=\int_\sigma^\tau n(r)\,dr, \tag{6.13}$$

where $n(r)$ is the number of points q in R^ϵ where $\rho(q)=r$.

For later reference, note that

$$\rho(q)\leqq b \text{ and } f(\rho(q),q)\geqq\epsilon \Rightarrow q\in R^\epsilon \tag{6.14}$$

In order to prove the theorem, it will be shown that if $\epsilon>0$, then there exists an $\eta=\eta(\epsilon;a,b)>0$ such that if $\delta_1,\cdots,\delta_k$ is a set of non-overlapping intervals on $a\leqq r\leqq b$, then

$$\sum_{i=1}^k|\delta_i H|<2\epsilon L_0 \text{ whenever } \sum_{i=1}^k|\delta_i|<\eta, \tag{6.15}$$

where $\delta_i H = H(\tau) - H(\sigma)$, $|\delta_i| = \tau - \sigma$ if $\delta_i = (\sigma, \tau]$.

Let $c = c(b)$ denote a constant such that

$$|f(p,q)| \leqq c \text{ for } 0 \leqq p \leqq b,\ 0 \leqq q \leqq L_0, \tag{6.16}$$

$$|f(p_2,q) - f(p_1,q)| \leqq c\,|p_2 - p_1| \text{ for } 0 \leqq p_1 \leqq p_2 \leqq b. \tag{6.17}$$

Note that, by (6.8),

$$\delta_i L = \int_0^{L_0} f(\sigma,q)\,\delta_i\chi\,dq + \int_0^{L_0} \chi(\tau,q)\,\delta_i f\,dq,$$

where, for example, $\delta_i\chi = \chi(\tau,q) - \chi(\sigma,q)$. Since $\delta_i\chi = 0$ unless $\sigma < \rho(q) \leqq \tau$, in which case $\delta_i\chi = -1$. Hence (6.17) shows that

$$\delta_i L = -\int_{S(i)} f(\rho(q),q)\,dq + 2\theta c L_0\,|\delta_i|,$$

where θ denotes a number (not always the same) satisfying $|\theta| \leqq 1$ and

$$S(i) = \{q : \sigma < \rho(q) \leqq \tau\} = \{q : \rho(q) \in \delta_i\}. \tag{6.18}$$

In view of (6.10), one has

$$\delta_i J = \sum_{\sigma < t \leqq \tau} \int_{Q(t)} f(t,q)\,dq.$$

Since $q \in Q(t)$ implies that $\rho(q) = t$,

$$\delta_i J = \int_{R(i)} f(\rho(q),q)\,dq$$

where

$$R(i) = \bigcup_{\sigma < t \leqq \tau} Q(t). \tag{6.19}$$

Clearly, $R(i) \subset S(i)$. Consequently, $H = L + J$ gives

$$|\delta_i H| \leqq \int_{S(i)-R(i)} f(\rho(q),q)\,dq + 2cL_0\,|\delta_i|. \tag{6.20}$$

In order to appraise the last integral, write $S(i) - R(i)$ as the union of three sets $S_1(i)$, $S_2(i)$, $S_3(i)$ defined as follows: Let $O(i)$ be an open set containing $R(i)$ and the discrete set of q-values in $S(i)$, if any, for which $z(\rho(q),q)$ is an isolated anormal point and such that $|O(i) - R(i)| < |\delta_i|$. Put

$$S_1 = [S(i) - R(i)] \cap O(i), \text{ so } |S_1| < |\delta_i|. \tag{6.21}$$

Define S_2, S_3 by

$$S_2 = [S(i) - R(i)] \cap \{q : f(\rho(q), q) < \epsilon\}, \tag{6.22}$$

$$S_3 = [S(i) - R(i)] - [S_1 \cup S_2]. \tag{6.23}$$

In particular, if $q \in S_3$, then $z(\rho(q), q)$ is normal and $d\rho/dq \neq 0$. The sets S_1, S_2 are not necessarily disjoint. Since $f(\rho(q), q) \geqq 0$,

$$\int_{S(i)-R(i)} \leqq \int_{S_1} + \int_{S_2} + \int_{S_3}. \tag{6.24}$$

By (6.16) and (6.21),

$$\int_{S_1} \leqq c\,|S_1| \leqq c\,|\delta_i|; \tag{6.25}$$

and, by (6.22),

$$\int_{S_2} \leqq \epsilon\,|S_2| \leqq \epsilon\,|S(i)|. \tag{6.26}$$

The closure of S_3 contains only points q such that $z(\rho(q), q)$ is normal and $d\rho/dq \neq 0$; cf. (6.21)-(6.23). Thus, $d\rho/dq \neq 0$ is continuous on the closure of S_3 and so, there exists a number $c_1(\epsilon) = c_1(\epsilon, a, b) > 0$ such that

$$q \in S_3 \Rightarrow |d\rho/dq| \geqq c_1(\epsilon) > 0. \tag{6.27}$$

Note that, by (6.14) and (6.22)-(6.23), $S_3 \subset R^\epsilon$. Thus, it is meaningful to speak of the total variation $T(S_3)$ of $\rho_b(q)$ over S_3. By (6.27),

$$c_1(\epsilon)\,|S_3| \leqq T(S_3)\text{; so that, by (6.13), } c_1(\epsilon)\,|S_3| \leqq \int_{\delta_i} n(r)\,dr.$$

Consequently, the last integral in (6.24) satisfies

$$\int_{S_3} \leqq c\,|S_3| \leqq (c/c_1(\epsilon)) \int_{\delta_i} n(r)\,dr. \tag{6.28}$$

In view of (6.20), (6.24), (6.25), (6.26), and (6.28),

$$\sum_{i=1}^{k} |\delta_i H| \leqq c(2L_0 + 1) \sum_{i=1}^{k} |\delta_i| + \epsilon L_0 + (c/c_1(\epsilon)) \sum_{i=1}^{k} \int_{\delta_i} n(r)\,dr.$$

Since the function $n(r)$, defined in connection with (6.9), is Lebesgue summable for $0 \leqq r \leqq b$, it follows that there exists an $\eta = \eta(\epsilon) > 0$ for which (6.15) holds. This proves Theorem 6.2.

Proof of Corollary 6.1. For $r \geqq 0$, let

$$j(r) = -[\phi(L(r)) - \phi(L(r-0))];$$

so that $j(r) \geqq 0$ and $j(r) = 0$ unless $r \geqq 0$ is a discontinuity point of $L(r)$ (i.e., of $J(r)$). Put

$$J_\phi(r) = \sum_{0 \leqq t \leqq r} j(t) \text{ for } r \geqq 0.$$

This is a finite sum or a convergent series for if $0 \leqq \phi'(L) \leqq c$ on a suitable L-interval, then

$$\sum_{0 \leqq t \leqq r} j(t) \leqq c \sum_{0 \leqq t \leqq r} [J(t) - J(t-0)] = cJ(r).$$

The function

$$H_\phi(r) = \phi(L(r)) + J_\phi(r)$$

is continuous. In fact, the uniform Lipschitz continuity of $\phi(L)$ on bounded L-intervals implies that $H_\phi(r)$ is absolutely continuous, since $L(r)$ has no purely singular component. Furthermore, for almost all r, $dH_\phi(r)/dr = \phi'(L(r))L'(r)$. Hence

$$H_\phi(b) - H_\phi(a) = \int_a^b \phi'(L(r))L'(r)\,dr \text{ for } 0 \leqq a < b < \infty.$$

Since $J_\phi(r)$ is non-decreasing, this gives the first part of (6.12). As the proof of the second part is similar, the proof of Corollary 6.1 is complete.

According as $\bar{r} \leqq r < 0$ or $r > 0$, let $D(r)$ denote the intersection of the (p,q)-set D in (6.6) and the strip $r \leqq p \leqq 0$ or the strip $0 \leqq p \leqq r$. Let $\Delta(r)$ denote the subset of M which is the image of $D(r)$ under the map $(p,q) \rightarrow z(p,q)$. Then, if r is non-exceptional, $\Delta(r)$ is the subset of M in (6.5) which is bounded by $\mathcal{B}$ and $\gamma(r)$. By (2.2),

$$(6.29) \qquad A(r) = \int\int_{\Delta(r)} dA = \pm \int_0^r \int_0^{L_0} f(p,q)\chi(p,q)\,dq\,dp \text{ if } \pm r > 0.$$

In particular, the area $A(\mathcal{B})$ enclosed by $\mathcal{B}$ is $A(\bar{r})$.

From (6.8) and (6.29),

$$(6.30) \qquad dA(r)/dr = \pm L(r) \text{ for almost all } r \geqq \bar{r}.$$

This leads to

COROLLARY 6.2. *If $A(\mathcal{B})$ is the (Riemann) area of the interior of $\mathcal{B}$, then*

$$(6.31) \qquad A(\mathcal{B}) = \int_{\bar{r}}^0 L(p)\,dp.$$

This result is given by Fiala [4], p. 332, under his assumptions. In particular, the results of the second part of Fiala's paper are valid under the conditions of this paper since they depend only on the inequalities (6.3)-(6.4), (6.11) and on (6.7), (6.30).

Remark. The part of Ahlfors paper [1], p. 5, depending on Fiala's results, uses (6.3) and the following inequality:

$$L^2(b) - L^2(a) \leqq 2 \int_a^b L(r)L'(r)\,dr \text{ for } 0 \leqq a \leqq b.$$

As this is a consequence of Corollary 6.1 with $\phi(L) = L^2$, Ahlfors' arguments remain correct under the conditions of this paper.

7. Isoperimetric inequalities. In addition to the assumptions in § 1 on M, assume that

$$\omega \equiv \int\int_M |K(x,y)|\,dA < \infty, \tag{7.1}$$

where $dA = (EG - F^2)^{\frac{1}{2}} dxdy$. Put

$$\mu = \int\int_M K(x,y)\,dA. \tag{7.2}$$

Let $\mathcal{B}$ be a rectifiable Jordan curve in N; $L(\mathcal{B})$ its (Riemann) length and $A(\mathcal{B})$ the (Riemann) area of the interior $\mathfrak{I}$ of $\mathcal{B}$. Then

$$L^2(\mathcal{B})/A(\mathcal{B}) \geqq 4\pi - 2\mu^+(\mathcal{B}), \tag{7.3}$$

where

$$\mu^+(\mathcal{B}) = \int\int_{\mathfrak{I}} \max(0, K(x,y))\,dA. \tag{7.4}$$

The isoperimetric inequality (7.3) in the analytic case is given by Fiala [4] who assumes $K(x,y) \geqq 0$ instead of (7.1). (In this case, the convergence of the integral in (7.1) follows from the assumption of completeness; Cohn-Vossen [3].) Fiala's proof of (7.3) is also valid under the more general condition (7.1). A different proof of (7.3) is given by Huber (cf. [8]) under relaxed conditions of differentiability.

In general, the following form of (7.3) is false

$$L^2(\mathcal{B})/A(\mathcal{B}) \geqq 4\pi - 2\mu(\mathcal{B}), \tag{7.5}$$

where

$$\mu(\mathcal{B}) = \int\int_{\mathfrak{I}} K(x,y)\,dA. \tag{7.6}$$

It will be shown, however, that (7.5) is "nearly" correct if $\mathcal{C}$ is replaced by the outermost component of $\gamma(r)$ for large non-exceptional $r > 0$.

THEOREM 7.1. *Let M satisfy the local smoothness and completeness assumptions of* § 1. *Assume also that* (7.1) *holds. Let $\mathcal{C}$ be a Jordan curve of class C^2. Let $r > 0$ be non-exceptional, $L^*(r)$ the length of the outermost component* (= *Jordan curve*) $\mathfrak{C}^*(r)$ *of the r-locus* $\gamma(r)$, *and $A^*(r)$ the area of the interior of $\mathfrak{C}^*(r)$. Then*

$$(7.7)\qquad \begin{aligned} L^{*2}(r)/A^*(r) &\to 4\pi - 2\mu \text{ as } r \to \infty, \\ L^2(r)/A(r) &\to 4\pi - 2\mu \text{ as } r \to \infty. \end{aligned}$$

This is contained in Fiala [4] for the analytic case when $K(x,y) \geqq 0$. Note that Theorem 7.1 implies

COROLLARY 7.1. *Let M be as in Theorem* 7.1. *Then*

$$(7.8)\qquad \mu \leqq 2\pi.$$

The statement that (7.8) follows from the completeness of M and (7.1) is due to Cohn-Vossen [3], under weaker assumptions of differentiability than in Theorem 7.1. For another proof, under still weaker conditions of differentiability, see Huber [7].

The proof of Theorem 7.1 will involve an isoperimetric inequality similar to (7.5) which depends on the position of $\mathcal{C}$ relative to the set of points $(x,y) \in \mathcal{J}$, where $|K(x,y)|$ is not small. Roughly speaking, it is required that these points be near the "center of $\mathcal{J}$."

LEMMA 7.1. *Let M be as above. Let $\mathcal{C}$ be a rectifiable Jordan curve; $\mathcal{J}$ its interior;* $-d(z) = \operatorname{dist}(z, \mathcal{C})$ *for* $z \in \mathcal{J}$; *and* $\bar{r} = \min d(z)$ *for* $z \in \mathcal{J}$. *Let* $\epsilon > 0$ *and let* p_ϵ *be chosen so that* $\bar{r} \leqq p_\epsilon < 0$ *and*

$$(7.9)\qquad \iint_{\mathcal{J}-\{z:\, d(z)\leqq p_\epsilon<0\}} |K(x,y)|\, dA \leqq \epsilon$$

and put

$$(7.10)\qquad \lambda = 1 - p_\epsilon/\bar{r}, \text{ so that } 0 \leqq \lambda < 1.$$

Then

$$(7.11)\qquad L^2(\mathcal{C})/2A(\mathcal{C}) \geqq 2\pi - \mu(\mathcal{C}) - \omega\lambda^2 - 2\epsilon.$$

Proof. It can be supposed that $\mathcal{C}$ is of class C^2, for otherwise $\mathcal{C}$ can be suitably approximated. The domain of integration in (7.9) is $\Delta(p_\epsilon)$ in the notation of Theorem 6.1. Hence

$$(7.12)\qquad |C(r)| \leqq \iint_{\Delta(r)} |K(x,y)|\, dA \leqq \epsilon \text{ for } 0 > r > p_\epsilon.$$

By the theorem of Gauss-Bonnet, the integral in (6.4) is $2\pi-\mu(\mathfrak{B})$. Hence (6.4) and (6.11) give

$$L(p)\leqq L(0)+[2\pi-\mu(\mathfrak{B})]p-\int_{p}^{0}C(r)\,dr \tag{7.13}$$

for $\bar{r}\leqq p<0$, where $L(0)=L_0=L(\mathfrak{B})$. Another integration and (6.31) give

$$A(\mathfrak{B})\leqq -L(0)\bar{r}-[2\pi-\mu(\mathfrak{B})]\bar{r}^2/2-\int_{\bar{r}}^{0}(\bar{r}-r)C(r)\,dr. \tag{7.14}$$

From $|C(r)|\leqq\omega$ and (7.12), it is seen that

$$\left|\int_{\bar{r}}^{0}(\bar{r}-r)C(r)\,dr\right|\leqq\omega\int_{\bar{r}}^{p_\epsilon}(r-\bar{r})\,dr+\epsilon p_\epsilon\bar{r}.$$

This is majorized by

$$\omega(p_\epsilon-\bar{r})^2/2+\epsilon\bar{r}^2\leqq(\bar{r}^2/2)[\omega\lambda^2+2\epsilon].$$

Hence, by (7.14),

$$A(\mathfrak{B})\leqq -L(0)\bar{r}-c\bar{r}^2/2,$$

where c is the number on the right of (7.11).

If $c\leqq 0$, then (7.11) is trivial. If $c>0$, the last inequality can be written as

$$A(\mathfrak{B})\leqq L^2(0)/2c-(1/2c)[L(0)-c\bar{r}]^2.$$

This gives (7.11).

Proof of Theorem 7.1. It will first be shown that the Lemma 7.1 implies that

$$\liminf_{r\to\infty}L^{*2}(r)/A^*(r)\geqq 4\pi-2\mu. \tag{7.15}$$

To this end, note that there exists an $r_\epsilon>0$ so large that if $\mathfrak{I}^*(r)$ is the interior of $\mathfrak{A}^*(r)$, then

$$\iint_{M-\mathfrak{I}^*(r)}|K(x,y)|\,dA<\epsilon \text{ if } r\geqq r_\epsilon. \tag{7.16}$$

Let $r>r_\epsilon$ and apply Lemma 7.1 with $\mathfrak{B}$ replaced by $\mathfrak{A}^*(r)$. The corresponding number $\bar{r}$ is a number r^* satisfying $-r^*\leqq r-\bar{r}$. Choose the analogue of p_ϵ to be $p_\epsilon^*=-(r-r_\epsilon)$. Since $z\in\mathfrak{I}^*(r_\epsilon)$ implies the analogue of $d(z) < p_\epsilon$, the inequality (7.16) shows that the analogue of (7.9) holds, so that p_ϵ^* is an admissible choice. It is also seen that

$$0 < \lambda \leqq (r_\epsilon - \bar{r})/(r - \bar{r}) \to 0 \text{ as } r \to \infty.$$

Hence (7.15) follows from (7.11).

Note that, by (6.3)

$$dL/dr \leqq 2\pi - \mu(\mathcal{B}) - C(r)$$

for non-exceptional $r > 0$. It is clear that

$$\mu(\mathcal{B}) + C(r) \to \mu \text{ as } r \to \infty.$$

Hence, for large non-exceptional $r > 0$,

$$dL(r)/dr \leqq 2\pi - \mu + \epsilon, \tag{7.17}$$

so that

$$2L(r)\,dL(r)/dr \leqq 2(2\pi - \mu + \epsilon)L(r).$$

Thus, Corollary 6.1 (cf. the Remark at the end of § 6) shows that if r_0 is sufficiently large,

$$L^2(r) - L^2(r_0) \leqq 2(2\pi - \mu + \epsilon)[A(r) - A(r_0)] \text{ for } r > r_0.$$

Hence

$$\limsup_{r\to\infty} L^2(r)/A(r) \leqq 4\pi - 2\mu. \tag{7.18}$$

Since $L^*(r) \leqq L(r)$ and $A^*(r) \geqq A(r)$, Theorem 7.1 follows from (7.15) and (7.18).

Remark. The proof can be modified to show that, under the conditions of Theorem 7.1,

$$\begin{aligned} L(r)/r &\to 2\pi - \mu \text{ as } r \to \infty, \\ L^*(r)/r &\to 2\pi - \mu \text{ as } r \to \infty, \end{aligned} \tag{7.19}$$

and so, by l'Hôpital's rule,

$$A(r)/(\tfrac{1}{2}r^2) \to 2\pi - \mu \text{ as } r \to \infty. \tag{7.20}$$

(This is also given by Fiala under his conditions of analyticity and $K(x,y) \geqq 0$).

In order to prove (7.19), note that under the assumptions of Lemma 7.1, one can deduce the inequality

$$L(\mathcal{B})/(-\bar{r}) \geqq 2\pi - \mu(\mathcal{B}) - \omega\lambda - \epsilon$$

from (7.13) with $p = \bar{r}$ and $L(0) = L(\mathcal{B})$. An application of this inequality, with $\mathcal{B}$ replaced by $\mathcal{A}^*(r)$, as in the proof of Theorem 7.1, gives

$$\liminf_{r \to \infty} L^*(r)/r \geqq 2\pi - \mu.$$

On the other hand,

$$\limsup_{r \to \infty} L(r)/r \leqq 2\pi - \mu$$

follows from (7.17) and Corollary 6.1. Since $L^*(r) \leqq L(r)$, this proves (7.19).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] L. V. Ahlfors, "Geodesic curvature and area," *Studies in Mathematical Analysis and Related Topics*, Stanford (1962).

[2] S. Banach, "Sur les lignes rectifiables et les surfaces dont l'aire est finie," *Fundamenta Mathematicae*, vol. 7 (1925), pp. 225-236.

[3] S. Cohn-Vossen, "Kürzeste Wege und Totalkrümmung auf Flächen," *Composito Mathematica*, vol. 2 (1935), pp. 69-133.

[4] F. Fiala, "Le problème des isopérimètres sur les surfaces ouvertes à courbure positive," *Commentarii Mathematici Helvetici*, vol. 13 (1940-41), pp. 293-346.

[5] P. Hartman, "On the local uniqueness of geodesics," *American Journal of Mathematics*, vol. 72 (1950), pp. 723-730.

[6] P. Hartman and A. Wintner, "On the third fundamental form of a surface," *ibid.*, vol. 75 (1953), pp. 298-334.

[7] A. Huber, "On subharmonic functions and differential geometry in the large," *Commentari Mathematici Helvetici*, vol. 32 (1957), pp. 13-72.

[8] ———, "Zur isoperimentrischen Ungleichung auf gekrümmten Flächen," *Acta Mathematica*, vol. 97 (1957), pp. 95-101.

[9] H. Hopf and W. Rinow, "Ueber der Begriff der vollständigen differentialgeometrischen Fläche," *Commentarii Mathematici Helvetici*, vol. 3 (1931), pp. 209-225.

[10] S. B. Meyers, "Connections between differential geometry, I. Simply connected surfaces," *Duke Mathematical Journal*, vol. 1 (1935), pp. 376-391.

[11] H. Poincaré, "Sur les lignes géodésiques de surfaces convexes," *Transactions of the American Mathematical Society*, vol. 6 (1905), pp. 237-274.

[12] J. H. C. Whitehead, "On the covering of a complete space by the geodesics through a point," *Annals of Mathematics*, vol. (2) 36 (1935), pp. 679-704.

ON RIEMANN'S METHOD AND A VARIATION BY MARTIN.

By A. G. Mackie.

1. Riemann and Martin functions. In 1860 Riemann [4] gave an account of a method of solving Cauchy's problem for a linear, second order, hyperbolic differential equation in two independent variables and this is now generally referred to as Riemann's method. If the differential operator L in its normal hyperbolic form is defined by

$$L[\phi] \equiv \frac{\partial^2 \phi}{\partial x \partial y} + g \frac{\partial \phi}{\partial x} + f \frac{\partial \phi}{\partial y} + c\phi, \tag{1}$$

where g, f and c are functions of x and y, then the solution of the equation

$$L[\phi] = F(x, y)$$

which, together with its normal derivative, takes prescribed values on some non-characteristic curve $\mathscr{L}$ can be given explicitly in terms of the Riemann function $R(x, y; x_0, y_0)$ associated with the operator. This is a function of the active variables x, y and the parameters x_0, y_0. As a function of x, y it satisfies the adjoint equation

$$L^*[R] \equiv \frac{\partial^2 R}{\partial x \partial y} - \frac{\partial}{\partial x}(gR) - \frac{\partial}{\partial y}(fR) + cR = 0. \tag{2}$$

Further, R obeys the relations

$$\frac{\partial R}{\partial x} - fR = 0 \quad \text{on } y = y_0, \tag{3}$$

$$\frac{\partial R}{\partial y} - gR = 0 \quad \text{on } x = x_0 \tag{4}$$

and the normalizing condition

$$R(x_0, y_0; x_0, y_0) = 1. \tag{5}$$

If the data are given on the curve $\mathscr{L}$, the essence of Riemann's method is to integrate the identity

$$\begin{aligned} &RL[\phi] - \phi L^*[R] \\ &= \tfrac{1}{2}\frac{\partial}{\partial x}\left(R\frac{\partial \phi}{\partial y} - \phi\frac{\partial R}{\partial y} + 2gR\phi\right) + \tfrac{1}{2}\frac{\partial}{\partial y}\left(R\frac{\partial \phi}{\partial x} - \phi\frac{\partial R}{\partial x} + 2fR\phi\right) \end{aligned}$$

Received April 7, 1964.
Revised June 15, 1964.

over the area PAB, where P is an arbitrary point with co-ordinates (x_0, y_0) and PA, PB are the x and y characteristics through P, A and B lying on $\mathscr{L}$, and then use Green's theorem. Because of equations (2), (3), (4) and (5), this gives, if $\mathscr{L}$ has negative slope,

$$\phi(x_0, y_0) = \tfrac{1}{2}(\phi R)_A + \tfrac{1}{2}(\phi R)_B \tag{6}$$
$$+\tfrac{1}{2}\int_A^B \left(R\frac{\partial\phi}{\partial x} - \phi\frac{\partial R}{\partial x} + 2fR\phi\right)dx - \left(R\frac{\partial\phi}{\partial y} - \phi\frac{\partial R}{\partial y} + 2gR\phi\right)dy$$
$$+ \iint\limits_{PAB} R(x, y; x_0, y_0)F(x, y)\,dx\,dy.$$

Minor sign changes are necessary if the curve $\mathscr{L}$ has positive slope. The same Riemann function can also be used to solve the characteristic boundary value problem wherein ϕ (but not its normal derivative) is given on two intersecting characteristics.

In 1943 Martin [3] described an alternative but similar method which is applicable to equations of the form

$$\frac{\partial^2\phi}{\partial x\partial y} + \alpha(x, y)\left(\frac{\partial\phi}{\partial x} + \frac{\partial\phi}{\partial y}\right) = F(x, y). \tag{7}$$

Although the differential operator in this equation is not as general as that in (1) it is nevertheless a very important special case as it includes many of the equations which are of most interest, including that which arises in one-dimensional gas dynamics and which prompted Riemann's original investigation. Martin's method also depends on the determination of an auxiliary function of x, y, x_0 and y_0 which he calls the resolvent but we find more convenient to call the Martin function. This function $M(x, y; x_0, y_0)$ must satisfy

$$\frac{\partial^2 M}{\partial x\partial y} - \alpha(x, y)\left(\frac{\partial M}{\partial x} + \frac{\partial M}{\partial y}\right) = 0 \tag{8}$$

(which is not, it should be noted, in general the adjoint equation) and the extra conditions

$$\frac{\partial M}{\partial x} = 1 \text{ on } y = y_0, \tag{9}$$

$$\frac{\partial M}{\partial y} = 1 \text{ on } x = x_0, \tag{10}$$

$$M(x_0, y_0; x_0, y_0) = 0. \tag{11}$$

These may be expressed in the form

$$(12) \qquad M(x, y_0; x_0, y_0) = x - x_0, \qquad M(x_0, y; x_0, y_0) = y - y_0.$$

In its applications the Martin function may be indeterminate to the extent of an arbitrary constant so that (9) and (10) suffice to define it. However, it is convenient to determine it uniquely by (11) for the sake of completeness.

If (7) is multiplied by $\dfrac{\partial M}{\partial x} + \dfrac{\partial M}{\partial y}$, (8) by $\dfrac{\partial \phi}{\partial x} + \dfrac{\partial \phi}{\partial y}$ and the resulting equations added, then

$$\frac{\partial}{\partial x}\left(\frac{\partial M}{\partial y}\frac{\partial \phi}{\partial y}\right) + \frac{\partial}{\partial y}\left(\frac{\partial M}{\partial x}\frac{\partial \phi}{\partial x}\right) = \left(\frac{\partial M}{\partial x} + \frac{\partial M}{\partial y}\right) F(x,y).$$

If this equation is integrated over PAB and Green's theorem is used,

$$(13) \qquad \begin{aligned} \phi(x_0, y_0) = \tfrac{1}{2}\phi_A + \tfrac{1}{2}\phi_B + \tfrac{1}{2}\int_A^B \frac{\partial M}{\partial x}\frac{\partial \phi}{\partial x}\,dx - \frac{\partial M}{\partial y}\frac{\partial \phi}{\partial y}\,dy \\ + \tfrac{1}{2}\iint\limits_{PAB}\left(\frac{\partial M}{\partial x} + \frac{\partial M}{\partial y}\right) F(x,y)\,dx\,dy. \end{aligned}$$

We have slightly generalized Martin's original result, which was established in [3] only for the homogeneous equation (7) with $F(x,y) = 0$, and for the following reason. The general similarity between the two methods suggests a close relationship between the Riemann and Martin functions. Yet if we set $F(x,y) = 0$ in (6) and (13) and compare the remaining terms, writing $f = g = \alpha$ in (6), there is no obvious simple way in which the functions R and M appear to be related. On the other hand, if we assume that $F(x,y)$ is non-zero but that the given values of ϕ and its normal derivative on $\mathscr{L}$ are zero, then all the terms on the right hand sides of (6) and (13) disappear except for the final double integral term in each case. Since $F(x,y)$ and the curve $\mathscr{L}$ are arbitrary, it follows that the Riemann and Martin functions are related by

$$(14) \qquad R = \tfrac{1}{2}\left(\frac{\partial M}{\partial x} + \frac{\partial M}{\partial y}\right).$$

Having established this remarkably simple relationship between the two functions, we should now verify that those parts of the solutions (6) and (13) which correspond to non-zero data on $\mathscr{L}$, that is the right hand sides without the double integral terms, are in fact identical when use is made of (14). This is not difficult. We first set $F = 0$, $f = g = \alpha$ in (6) and then substitute for R in terms of M by means of (14). The function $\alpha(x,y)$ is now eliminated by using (8) since it does not occur explicitly in (13). There results a line integral along AB together with some additional terms and when use is made of relations such as

$$\left(\quad \frac{\partial M}{\partial y}\right)_B - \left(\phi\,\frac{\partial M}{\partial y}\right)_A = \int_A^B \frac{\partial}{\partial x}\left(\phi\,\frac{\partial M}{\partial y}\right)\,dx + \frac{\partial}{\partial y}\left(\phi\,\frac{\partial M}{\partial y}\right)\,dy,$$

equation (13) is recovered. We omit the details.

Equation (14) shows how the Riemann function can be easily obtained when the Martin function is known. The process can be reversed for from (8) and (14)

$$\frac{\partial^2 M}{\partial x \partial y} = 2\alpha(x, y)R.$$

Hence, using (9), (10) and (11),

$$\begin{aligned} (15) \quad & M(x, y\,; x_0, y_0) \\ & = 2\int_{x_0}^{x}\int_{y_0}^{y} \alpha(X, Y)R(X, Y\,; x_0, y_0)\,dX\,dY + x + y - x_0 - y_0. \end{aligned}$$

Thus any method by which Riemann functions can be obtained is also applicable, because of equation (15), to the problem of finding Martin functions; and conversely. Copson [1] has made a survey of the known methods of obtaining Riemann functions. As a result of (15), this survey can be regarded as covering the Martin functions for any equations included in it which are of the form (7).

2. The Euler-Poisson equation. We illustrate this equivalence principle by considering the example discussed by Martin himself in which $\alpha(x, y) = N/(x+y)$ and there results the Euler-Poisson equation

$$\frac{\partial^2 \phi}{\partial x \partial y} + \frac{N}{x+y}\left(\frac{\partial \phi}{\partial x} + \frac{\partial \phi}{\partial y}\right) = 0,$$

where N is a constant. The Riemann function for this equation can be expressed in terms of a hypergeometric function or a Legendre function (see [1]) and the Martin function is given in [3] in terms of Appell's hypergeometric functions in two variables. It is not easy to deduce one function from the other by (14) or (15) from these representations because of the complicated way in which they contain x and y. Some years ago the author, in an unpublished note, suggested a representation of the Riemann function as a contour integral and this was included by Copson [1] in his survey. We shall not discuss the motivation for this approach here as this is fully explained in [1] but merely quote the result which is

$$(16)\qquad R(x,y;x_0,y_0)=\frac{x+y}{2\pi i}\int_{C_1}\frac{(\zeta-x)^{N-1}(\zeta+y)^{N-1}}{(\zeta-x_0)^N(\zeta+y_0)^N}\,d\zeta.$$

For non-integral values of N the integrand has branch points at $\zeta=x$, x_0, $-y$, $-y_0$. (For convenience we may assume that x, x_0, y, y_0 are all positive.) If those connecting x, x_0 and $-y$, $-y_0$ are joined by cuts, then the integrand is analytic and single-valued in the ζ-plane cut in this way. The contour C_1 surrounds the right hand cut and for what follows it is convenient to introduce a second contour C_2 surrounding the left hand cut in the opposite direction. Since the integrand in (16) is of order ζ^{-2} for large $|\zeta|$, we can deform C_1 into the integral from $i\infty$ to $-i\infty$ and from there into C_2.

Equation (16) is clearly in a suitable form to apply (15) as the term $\alpha(x,y)$ cancels the factor outside the integral. Thus

$$\begin{aligned}M(x,y;x_0,y_0)&\\ =-\frac{1}{\pi i N}\int_{C_1}&\frac{\{(\zeta-x)^N-(\zeta-x_0)^N\}\{(\zeta+y)^N-(\zeta+y_0)^N\}}{(\zeta-x_0)^N(\zeta+y_0)^N}\,d\zeta\\ &+x+y-x_0-y_0.\end{aligned}$$

By the argument of the preceding paragraph this may also be written with C_2 replacing C_1, and if the mean of the two is taken, then

$$(17)\qquad M(x,y;x_0,y_0)=-\frac{1}{2\pi i N}\int_{C_1+C_2}\frac{(\zeta-x)^N(\zeta+y)^N}{(\zeta-x_0)^N(\zeta+y_0)^N}\,d\zeta,$$

since

$$(18)\qquad \frac{1}{2\pi i}\int_{C_1}\frac{(\zeta-x)^N}{(\zeta-x_0)^N}\,d\zeta=N(x_0-x)$$

with a similar result for the other integral involving y and y_0 round C_2. To prove (18) we first assume that $-1<N<1$, flatten the contour onto the real axis in order to evaluate it and then make appeal to the theory of analytic continuation to establish it for general values of N.

The contour in (17) can be regarded if desired as a figure of eight around the two branch cuts. When N is integral the Martin function is a rational function of x, y, x_0 and y_0 since the branch cuts now disappear and the integral in (17) can be evaluated by determining residues at multiple poles. The poles are at x_0, $-y_0$ when $N>0$ and at x, $-y$ when $N<0$. When $N=0$ we revert to (15), set $\alpha(X,Y)=0$ and obtain

$$(19)\qquad M(x,y;x_0,y_0)=x+y-x_0-y_0.$$

This is the case of the ordinary wave equation with Riemann function unity as verified by (14) and (19).

Although equation (17) was obtained by applying (15) to the contour integral representation of the Riemann function, it is interesting to comment briefly on how it should have been derived directly from the defining equations (12). It was shown by Martin that $M(x,y;x_0,y_0)$, while satisfying (8) as a function of x and y, satisfies the original homogeneous equation ((7) with $F(x,y)=0$) as a function of x_0 and y_0. (This property is shared by the Riemann function and the proof is similar.) Now $(\zeta-x)^N(\zeta+y)^N$ is a solution of (8) for an arbitrary value of ζ. Thus we are tempted to try an expression of the form

$$M(x,y;x_0,y_0)=\int_{C_1}\frac{(\zeta-x)^N(\zeta+y)^N}{(\zeta-x_0)^N(\zeta+y_0)^N}\,h_1(\zeta)\,d\zeta + \int_{C_2}\frac{(\zeta-x)^N(\zeta+y)^N}{(\zeta-x_0)^N(\zeta+y_0)^N}\,h_2(\zeta)\,d\zeta,$$

since this will satisfy the correct equations as a function both of x, y and of x_0, y_0 for arbitrary analytic functions $h_1(\zeta)$, $h_2(\zeta)$. When $y=y_0$ the integral round C_2 vanishes and to satisfy $M=x-x_0$ in the case we must solve

$$\int_{C_1}\frac{(\zeta-x)^N}{(\zeta-x_0)^N}\,h_1(\zeta)\,d\zeta=x-x_0$$

for $h_1(\zeta)$. The solution of such a contour integral equation with an arbitrary right hand side is given in [2], § 2 and using this we find that $h_1(\zeta)=-1/2\pi iN$. Similarly $h_2(\zeta)=-1/2\pi iN$ and equation (17) then follows as before.

3. Conclusions. In the preceding sections we have demonstrated the close connection between the methods of Riemann and of Martin for solving linear, second order, hyperbolic equations and between the two auxiliary functions upon which these methods depend. Two aspects would seem to be of some interest. The first is that, although in practice most of the applications are to homogeneous equations with non-homogeneous boundary conditions, much more insight is gained by considering the non-homogeneous equation with homogeneous boundary conditions. This, of course, is linked with the basic definition of the ordinary Green's function $G(x,y;x_0,y_0)$ which satisfies $L[G]=\delta(x-x_0,y-y_0)$ together with homogeneous boundary conditions although it is used to solve $L[\phi]=0$ where ϕ satisfies non-homogeneous boundary conditions. The second point of interest is the use of contour integrals which in this case enables the algebraic manipulation to be carried out with the minimum of effort.

In his original paper Martin states that the conditions defining the resolvent or Martin function are simpler than those for the Riemann function and that the form of solution as a line integral is also simpler. While this is certainly true it seems that some of the advantages to be gained in this way are offset by the fact that in general the Martin function is a more complicated expression than the Riemann function as might be expected from (15).

VICTORIA UNIVERSITY OF WELLINGTON,
NEW ZEALAND.

REFERENCES.

[1] E. T. Copson, "On the Riemann-Green function," *Archive for Rational Mechanics and Analysis*, vol. 1 (1958), pp. 324-348.

[2] A. G. Mackie, "Contour integral solutions of a class of differential equations," *Journal of Rational Mechanics and Analysis*, vol. 4 (1955), pp. 733-750.

[3] M. H. Martin, "The rectilinear motion of a gas," *American Journal of Mathematics*, vol. 65 (1943), pp. 391-407.

[4] B. Riemann, "Ueber die Fortpflanzung ebener Luftwellen von endlicher Schwingungsweite," *Abhandlungen der K. Gesellschaft der Wissenschaften zu Göttingen*, vol. 8 (1860).

NON-LOCAL ELLIPTIC BOUNDARY VALUE PROBLEMS.

By Felix E. Browder.*

Introduction. In recent years, an intensive investigation has been carried through by various writers of local or differential boundary value problems for linear elliptic differential operators A of order m which are *elliptic* or *coercive* in the sense that for u satisfying the homogeneous boundary conditions, there exists c independent of u such that:

$$\sum_{|\rho|\leqq m} \| D^\rho u \|_{L^2} \leqq c\{\| u \|_{L^2} + \| Au \|_{L^2}\} \tag{1}$$

(Cf., e.g., [1], [6]-[9], [12], [17] Chap. X, [18]-[21], [23]-[25]). This investigation has also been extended recently ([2]) to semi-local boundary value problems, i.e. with the boundary conditions determined by differential operators with singular integral operators as coefficients.

It is our object in the present paper to establish some useful general results concerning the general class of non-local elliptic boundary value problems for an elliptic operator A on a bounded domain G.

Let $W^{m,2}(G)$ be the Hilbert space $\{u \,|\, D^\alpha u \in L^2(G), |\alpha| \leqq m\}$ with inner product $(u,v)_m = \sum_{|\alpha|\leqq m} (D^\alpha u, D^\alpha v)$. Then by a general linear elliptic boundary value operator, we shall mean a bounded linear mapping γ of $W^{m,2}(G)$ onto a Hilbert space Γ such that the set $\{u \,|\, \gamma u = 0\}$ is dense in $L^2(G)$ and such that there exists $c > 0$ for which

$$\| u \|_{W^{m,2}(G)} \leqq c\{\| u \|_{L^2(G)} + \| Au \|_{L^2(G)}\} \tag{2}$$

for all u in $W^{m,2}(G)$ for which $\gamma u = 0$.

We can associate with each such boundary value problem $\{A, \Gamma, \gamma\}$ two linear operators which we denote by (A, γ) and A_γ. The first, (A, γ), is the bounded linear mapping of $W^{m,2}(G)$ into $L^2(G) \oplus \Gamma$ given by

$$(A, \gamma)(u) = [Au, \gamma u].$$

The second, A_γ, is the closed densely defined linear operator in $L^2(G)$ given by

$$D(A_\gamma) = \{u \,|\, u \in W^{m,2}(G), \gamma u = 0\}$$

Received April 15, 1964.

* Supported in part by the Sloan Foundation, the Army Research Office (Durham) through grant ARO(D)-31-124-G455, and National Foundation Science grants G-19751 and GP2283.

and

$$A_\gamma u = Au, \ u \in D(A_\gamma).$$

It is an easy consequence of the definition of ellipticity for $\{A, \Gamma, \gamma\}$ that each of these operators has a closed range and a finite-dimensional nullspace. There are two other types of analytical facts that are of especial interest in connection with such operators and whose proof for differential boundary value problems is not so direct. These are embodied in the following questions:

(1) *Are* (A, γ) *and* A_γ *Fredholm operators, i. e. do their ranges have finite codimension in the respective range spaces?*

(2) *Suppose* $\{A, \Gamma, \gamma\}$ *and* $\{A', \Gamma', \gamma'\}$ *are elliptic problems which are formally adjoint, i. e. such that*

$$(Au, v) = (u, A'v)$$

if $\gamma u = 0$, $\gamma' v = 0$. *Then, is it true that*

$$A'_{\gamma'} = (A_\gamma)^*$$

in the sense of adjoints of unbounded operators in Hilbert space? Is it true that the range of A_γ *coincides with the annihilator of the nullspace of* $A'_{\gamma'}$*?*

We construct here a general principle to answer such questions which we apply under detailed hypotheses elsewhere. To state this principle, we introduce the concept of a *continuous family* of elliptic boundary value problems: If Ω is a topological space and for each ω in Ω, $\{A(\omega), \Gamma, \gamma(\omega)\}$ is an elliptic boundary value problem of order m on G, the family is said to be continuous if the mapping $\omega \to (A(\omega), \gamma(\omega))$ is a continuous mapping of Ω into $L(W^{m,2}(G), L^2(G) \oplus \Gamma)$ with the uniform topology. Then our basic result is the following:

Let Ω *be a connected space,* $\{A(\omega), \Gamma, \gamma(\omega)\}$ *and* $\{A'(\omega), \Gamma', \gamma'(\omega)\}$ *two continuous families of elliptic boundary value problems which are formally adjoint for each* ω *in* Ω. *Suppose that for a single parameter value* ω_0, *we know that*

$$(A'(\omega_0))_{\gamma'(\omega_0)} = (A_0(\omega_0)_{\gamma(\omega_0)})^*.$$

Then for all ω *in* Ω,

$$(A'(\omega))_{\gamma'(\omega)} = (A(\omega)_{\gamma(\omega)})^*$$

while

$$(A(\omega), \gamma(\omega)) \text{ and } (A(\omega))_{\gamma(\omega)}$$

are Fredholm operators for every ω, *they have the same index, and this index is independent of* ω *in* Ω.

The practical import of this result is that if we can deform a given elliptic problem into a special problem $(A(w_0), \gamma(\omega_0))$ whose adjoint problem we can handle (in particular, a differential boundary value problem) while simultaneously deforming the formal adjoint into the adjoint problem, then we have shown that the formal adjoint is the real adjoint and in particular that (A, γ) is a Fredholm operator.

General boundary value problems for elliptic operators were first systematically studied for second-order operators by Calkin [13] and Visik [27]. If A is an elliptic operator of order m, A' its formal adjoint, Visik introduced the minimal and maximal operators of A:

A_0, the minimal operator of A, is the closure as an operator in $L^2(G)$ of A restricted to $C_0^\infty(G)$, the C^∞ functions with compact support in G;

A_1, the maximal operator of A, is given by

$$\begin{cases} D(A_1) = \{u.\mid u \in L^2(G); Au \in L^2(G) \text{ where } A \text{ is} \\ \text{taken as a distribution derivative}\} \\ A_1 u = Au. \end{cases}$$

Then the closed operator T is said to be the realization of A under a homogeneous boundary condition if

$$A_0 \subseteq T \subseteq A_1.$$

If we assume our boundary value operator γ is proper, i.e. $\gamma u = 0$ for $u \in C_0^\infty(G)$, then it is easy to see that A_γ is a realization of A in the above sense. Moreover the realization $T = A_\gamma$ for some elliptic boundary value problem $\{A, \Gamma, \gamma\}$ if and only if $D(T) \subset W^{m,2}(G)$. Thus the class of boundary value problems which we consider here contains the subclass of Visik's problems for which $D(T) \subset W^{m,2}(G)$.

Abstract boundary value problems have been studied for other purposes in Hörmander [16], Browder [9], [10], [11], Peetre [20] and Phillips [22]. An investigation of special classes of non-local elliptic boundary value problems along the lines of Calkin [13] has been made for second-order equations by Bade and Freeman [3] and Freeman [14] and for higher order equations by Beals [5].

In Section 1, we establish Theorem 1 which is an abstract analogue of our result on elliptic problems. In Section 2, we specialize the result of Theorem 1 to obtain our basic result in Theorem 2. The proof of Theorem 1 combines ideas from the writer's paper [8], [10] and [11] with an extremely simple but ingenious remark made by Beals in [4].

Section 1. Let H be a Hilbert space with inner product $(.,.)$, W another Hilbert space which is identified with a dense linear subset of H so that the injection map of W into H is compact.

Definition (1.1). *By an abstract elliptic boundary value problem for the pair* $\{H, W\}$ *is meant the following: We give a Hilbert space* Γ *and continuous linear mappings* A *of* W *into* H *and* γ *of* W *onto* Γ *such that for some constant* c

$$\| u \|_W \leqq c\{\| Au \|_H + \| u \|_H + \| \gamma u \|_\Gamma\} \tag{1.1}$$

for all u *in* W, *while* $\gamma = 0$ *on a dense subset of* H.

Corresponding to each abstract elliptic boundary value problem in the sense of Definition (1.1), we may form two operators as follows:

Definition (1.2). (a) (A,γ) *is the bounded linear mapping of* W *into* $H \oplus \Gamma$ *given by*

$$(A,\gamma)(u) = [Au, \gamma u].$$

(b) A_γ *is the densely defined linear operator in* H *with*

$$D(A_\gamma) = \{u \mid u \in W, \gamma u = 0\}$$
$$A_\gamma u = Au, u \in D(A_\gamma).$$

LEMMA (1.1). *If* $\{\Gamma, A_\gamma\}$ *defines an abstract elliptic boundary value problem for* $\{H, W\}$, *then* A_γ *is a closed linear operator in* H.

Proof of Lemma (1.1). For $\gamma u = 0$, the inequality (1.1) implies

$$\| u \|_W \leqq c\{\| A_\gamma u \|_H + \| u \|_H\} \tag{1.2}$$

for $u \in D(A_\gamma)$. Hence if $u_k \in D(A_\gamma)$, $u_k \to u$ in H, $A_\gamma u_k \to v$, it follows that $u_k \to u$ in W, $u \in W$, $\gamma u_k \to \gamma u$, $Au_k \to Au$: Hence $\gamma u = 0$, $v = Au$, i.e. $u \in D(A_\gamma)$, $A_\gamma u = v$. Thus A_γ is closed. Q.E.D.

Definition (1.3). *Let* $\{\Gamma, A, \gamma\}$ *and* $\{\Gamma', A', \gamma'\}$ *be two abstract elliptic boundary value problems for* $\{H, W\}$. *Then we say that these two boundary value problems are formally adjoint to one another if*

$$(A_\gamma u, v) = (u, A'_{\gamma'} v)$$

for all $u \in D(A_\gamma)$, $v \in D(A'_{\gamma'})$.

Definition (1.4). *Let Ω be a topological space, Γ a given Hilbert space, and for each ω in Ω, suppose we are given an abstract boundary value problem $\{\Gamma, A(\omega), \gamma(\omega)\}$ for $\{H, W\}$. Then this family of boundary value problems is said to be continuous if the mapping*

$$\omega \rightarrow (A(\omega), \gamma(\omega))$$

is a continuous mapping of Ω into $L(W, H \oplus \Gamma)$, where the latter has the norm topology.

LEMMA (1,2). *If $\{\Gamma, A, \gamma\}$ defines an abstract elliptic boundary value problem for $\{H, W\}$, then for each of the operators (A, γ) and A_γ of Definition* (1.2), *the nullspace is of finite dimension and the range is closed. The nullspace of A_γ coincides with the nullspace of (A, γ).*

Proof of Lemma (1.2). We note first that if $A_\gamma u = 0$, then $\gamma u = 0$ and $Au = 0$, i.e. $u \in N((A, \gamma))$ and conversely. For such u, we have by the inequality (1.1)

$$\| u \|_W \leqq c \| u \|_H.$$

Since the injection map of W into H is compact, the unit ball in $N(A_\gamma)$ is pre-compact. Hence $N(A_\gamma)$ is of finite dimension.

To prove that $R(A_\gamma)$ is closed in H, we remark that for $u \in D(A_\gamma)$,

$$\| u \|_W \leqq c\{\| u \|_H + \| A_\gamma u \|_H\}.$$

Hence the mapping J of the graph of A_γ into H given by $J([u, A_\gamma u]) = u$ is compact. It follows by Lemma 1 of [9] that the range of A_γ is closed in H. The same argument applies to (A, γ). Q. E. D.

Our basic result is the following:

THEOREM 1. *Let Ω be a connected topological space, $\{\Gamma, A(\omega), \gamma(\omega)\}$ and $\{\Gamma', A'(\omega), \gamma'(\omega)\}$ two continuous families of abstract elliptic boundary value problems for $\{H, W\}$. Suppose that:*

1) *For each ω in Ω, $\{\Gamma, A(\omega), \gamma(\omega)\}$ is formally adjoint to $\{\Gamma', A'(\omega), \gamma'(\omega)\}$.*

2) *For a particular ω_0 in Ω,*

$$(A'(\omega_0))_{\gamma'(\omega_0)} = ((A(\omega_0))_{\gamma(\omega_0)})^*.$$

Then:

a) *For every ω in Ω,*

$$(A'(\omega))_{\gamma'(\omega)} = ((A(\omega))_{\gamma(\omega)})^*.$$

b) *For every* ω *in* Ω, $(A(\omega),\gamma(\omega))$ *and* $(A'(\omega),\gamma'(\omega))$ *are Fredholm operators whose index does not depend on* ω *in* Ω.

c) *For each* ω *in* Ω, $(A(\omega))_{\gamma(\omega)}$ *and* $(A'(\omega))_{\gamma'(\omega)}$ *are Fredholm operators. Each index does not depend on* ω *in* Ω.

$$\textit{Index}\,(A(\omega),\gamma(\omega)) = \textit{index}\,(A(\omega))_{\gamma(\omega)},\quad \omega\in\Omega.$$

We shall obtain the proof of Theorem 1 as the culmination of a series of auxiliary results. We begin with the following:

LEMMA (1.3). *Under the hypotheses of Theorem* 1, *the conclusion* (a) *implies the conclusions* (b) *and* (c).

Proof of Lemma (1.3). If conclusion (a) holds, we have

$$(A'(\omega))_{\gamma'(\omega)} = (A(\omega)_{\gamma(\omega)})^{*}.$$

By Lemma (1.2), $(A(\omega))_{\gamma(\omega)}$ has a closed range and a finite-dimensional nullspace, as does $(A'(\omega))_{\gamma'(\omega)}$. Since the range of $(A(\omega))_{\gamma(\omega)}$ is closed in H, its codimension equals the dimension of $N((A(\omega)_{\gamma(\omega)})^{*})$, i.e. of $N((A'(\omega))_{\gamma'(\omega)})$, which is finite. Hence $(A(\omega))_{\gamma(\omega)}$ and $(A'(\omega))_{\gamma'(\omega)}$ are Fredholm operators.

As we have already remarked in Lemma (1.2),

$$N((A(\omega))_{\gamma(\omega)}) = N((A(\omega),\gamma(\omega))).$$

We know moreover that $\operatorname{codim} R(A_\gamma) = \operatorname{codim} R((A,\gamma))$ for any abstract elliptic boundary value problem. Indeed, since γ maps W onto Γ, there exists a bounded linear map S of Γ into W such that $\gamma S =$ the identity on Γ. Then $(A,\gamma)u = [v,f]\in H\oplus\Gamma$ if and only if

$$(A,\gamma)(u-Sf) = [v-ASf, 0],$$

i. e.

$$A_\gamma(u-Sf) = v-ASf.$$

Now suppose $\operatorname{codim} R(A_\gamma) = r$. Then there exist r elements $\{w_1, w_2,\cdots, w_r\}$ such that $w\in R(A_\gamma)$ if and only if $(w,w_j)=0$ for $1\leqq j\leqq r$. By the above argument, $[v,f]$ lies in $R((A,\gamma))$ if and only if $v-ASf$ lies in $R(A_\gamma)$ i.e. if and only if

$$(v-ASf,\omega_j) = 0, \qquad 1\leqq j\leqq r.$$

Hence $\operatorname{codim} R((A,\gamma))\leqq r = \operatorname{codim} R(A_\gamma)$. On the other hand, $w\in R(A_\gamma)$ if and only if $[w,0]\in R((A,\gamma))$. Hence $\operatorname{codim} R(A_\gamma)\leqq \operatorname{codim} R((A_\gamma))$, and finally $\operatorname{codim} R(A_\gamma) = \operatorname{codim} R(A_\gamma)$.

It follows from the conclusion of the preceding paragraph that if $(A(\omega))_{\gamma(\omega)}$ is a Fredholm operator, so is $(A(\omega),\gamma(\omega))$ and their indexes are equal. On the other hand, for ω near ω_1 and for a given $\epsilon>0$

$$\begin{aligned}\|(A(\omega),\gamma(\omega))(u)-(A(\omega_1),\gamma(\omega_1))(u)\|_{H\oplus\Gamma}\\ \leqq\epsilon\|u\|_W\leqq c\epsilon\{\|u\|_H+\|(A(\omega_1),\gamma(\omega_1))(u)\|_{H\oplus\Gamma}\}.\end{aligned}$$

If we choose ϵ sufficiently small, it follows by a standard theorem on invariance of the index under perturbations small in the graph norm ([15]) that index $(A(\omega),\gamma(\omega))=\text{index}(A(\omega_1),\gamma(\omega_1))$. Hence the index of $(A(\omega),\gamma(\omega))$ is locally constant in ω and since Ω is connected, the index of $(A(\omega),\gamma(\omega))$ is constant. Thus index $(A(\omega))_{\gamma(\omega)}=\text{index}\,(A(\omega),\gamma(\omega))=\text{constant}$. If we apply the same argument to the adjoint problem $(A'(\omega),\gamma'(\omega))$, we complete the proof of the lemma. Q. E. D.

We now state and prove a simple but ingenious lemma due to Richard Beals ([4]).

LEMMA (1.4). *Let T and T_1 be two closed Fredholm operators in H with $T\subseteq T_1$. Suppose that* $\text{index}\,T=\text{index}\,T_1$. *Then* $T=T_1$.

Proof of Lemma (1.4). If $T\subset T_1$, then $N(T)\subset N(T_1)$, $R(T)\subseteq R(T_1)$. Hence $\text{index}\,T\leqq\text{index}\,T_1$ and equality holds only if $N(T)=N(T_1)$, $R(T)=R(T_1)$. Then, however, $T=T_1$.

LEMMA (1.5). *Under the hypotheses of Theorem* 1, *the set*

$$\Omega_1=\{\omega\mid\omega\in\Omega,\ (A'(\omega))_{\gamma'(\omega)}=(A(\omega)_{\gamma(\omega)})^*\}$$

is open in Ω.

Proof of Lemma (1.5). For all ω in Ω, we have

$$(A'(\omega))_{\gamma'(\omega)}\subseteq(A(\omega)_{\gamma(\omega)})^*.$$

If $\omega_1\in\Omega_1$, then by the argument of the proof of Lemma (1.3), there exists a neighborhood N of ω_1 in Ω such that for ω in N, $(A(\omega))_{\gamma(\omega)}$ and $(A'(\omega))_{\gamma'(\omega)}$ are Fredholm operators and their indexes are constant. Then on N

$$\begin{aligned}\text{index}(A'(\omega))_{\gamma'(\omega)}-\text{index}((A(\omega)_{\gamma(\omega)})^*)\\ =\text{index}(A'(\omega))_{\gamma'(\omega)}+\text{index}((A(\omega))_{\gamma(\omega)})\end{aligned}$$

is constant and equals 0 at ω_1. Thus

$$\text{index}(A'(\omega)_{\gamma'(\omega)})=\text{index}((A\omega)_{\gamma(\omega)})^*)$$

for $\omega \in N$, so that by Lemma (1.4)

$$(A'(\omega))_{\gamma'(\omega)} = (A(\omega)_{\gamma(\omega)})^*$$

for $\omega \in N$. Hence $N \subset \Omega_1$. Q. E. D.

LEMMA (1.6). *Let $\omega \rightarrow \gamma(\omega)$ be a continuous mapping of Ω into $L(W, \Gamma)$ with the norm topology such that $\gamma(\omega)$ maps onto Γ for each ω. Then there exists a continuous map S of Ω into $L(\Gamma, w)$ with the norm topology such that*

$$\gamma(\omega) S_\omega = \text{identity on } \Gamma$$

for each ω in Ω.

Proof of Lemma (1.6). For each ω in Ω, let $\gamma^*(\omega)$ be the map of Γ into W adjoint to $\gamma(\omega)$. Since $\gamma(\omega)$ maps W onto Γ, $\gamma(\omega)^*$ is one-to-one, has closed range in W, and has a bounded inverse defined on its range. Let P_ω be the orthogonal projection operator of W in $R(\gamma(\omega)^*)$. If $\omega_1 \in \Omega$, the mapping of $R(\gamma(\omega_1)^*)$ onto $R(\gamma(\omega)^*)$ given by

$$K_\omega(u) = \gamma(\omega)^*(\gamma(\omega_1)^*)^{-1}(u)$$

maps $R(\gamma(\omega_1)^*)$ one-to-one onto $R(\gamma(\omega)^*)$, is norm-continuous in ω on Ω, and reduces to the identity map for $\omega = \omega_1$.

We now assert that P_ω is norm-continuous in ω on Ω. It suffices to establish this continuity near ω_1 for a given but arbitrary ω_1 in Ω. Let $u \in W$. Then $P_\omega u$ lies in $R(\gamma(\omega)^*)$ and by the remarks of the preceding paragraph, there exists an unique u_ω in $R(\gamma(\omega_1)^*)$ such that

$$K_\omega u_\omega = P_\omega u.$$

To show that P_ω is norm continuous at ω, it suffices to show for $\epsilon > 0$ that there exists a neghborhood N of ω_1 in Ω such that for ω in N, $\| u_\omega - u_{\omega_1} \|_W < \epsilon$ for all u in W with $\| u \| \leqq 1$. We know however that $u_{\omega_1} = P_{\omega_1} u$ and that for $\omega \in \Omega$,

$$u - K_\omega u_\omega \in R(\gamma(\omega)^*)^\perp$$

i. e.

$$(u - K_\omega u_\omega, K_\omega v) = 0 \tag{1.3}$$

for all v in $R(\gamma(\omega_1)^*)$. Let J be the injection map of $R(\gamma(\omega_1)^*)$ into W, so that J^* mapping W into $R(\gamma(\omega_1)^*)$ coincides with the orthogonal projection P_{ω_1}. Since K_ω is norm-continuous, $K_\omega = J + R_\omega$, where $\| R_\omega \| \rightarrow 0$ as $\omega \rightarrow \omega_1$. Hence $K_\omega^* = P_{\omega_1} + R_\omega^*$, where $\| R_\omega^* \| = \| R_\omega \| \rightarrow 0$ as $\omega \rightarrow \omega_1$.

The equality (1.3) holds for all v in $R(\gamma(\omega_1)^*)$ if and only if K_ω^* $(K_\omega u_\omega) = K_\omega^* u$, i. e. if and only if

$$u_\omega + P_{\omega_1} R_\omega u_\omega + R_\omega^* J u_\omega + R_\omega^* R_\omega u_\omega = P_{\omega_1} u + R_\omega^* u$$

We remark that if $\| R_\omega \| = \| R_\omega^* \| < \delta$, then

$$\| u_\omega \| \leqq \| u \| \{1 + \delta\} + (2\delta + \delta^2) \| u_\omega \|$$

so that choosing δ sufficiently small we have

$$\| u_\omega \| \leqq 2 \| u \| \leqq 2$$

for $\| u \| \leqq 1$. Then

$$\begin{aligned} \| u_\omega - u_{\omega_1} \| &= \| u_\omega - P_{\omega_1} u \| \\ &\leqq \| R_\omega \| + 2 \| R_\omega^* \| + 2 \| R_\omega \| \cdot \| R_\omega^* \| + 2 \| R_\omega^* \| \\ &\leqq 5\delta + 2\delta^2. \end{aligned}$$

If we choose $\delta > 0$ so small that $5\delta + 2\delta < \epsilon$, we have verified that u_ω is continuous in ω on Ω uniformly for all u with $\| u \| \leqq 1$.

For $\omega \in \Omega$, define $T_\omega \colon W \to \Gamma$ by $T_\omega u = (\gamma(\omega)^*)^{-1}(P_\omega u)$. Then $T_\omega u$ is well-defined for each u, and if $u = \gamma(\omega)^* v$, then

$$T_\omega(\gamma(\omega)^* v) = (\gamma(\omega)^*)^{-1}(\gamma(\omega)^* v) = v,$$

i. e.

$$T_\omega \gamma(\omega)^* = \text{identity on } \Gamma.$$

However,

$$\gamma(\omega)^* \{ (\gamma(\omega_1)^*)^{-1} u_\omega \} = P_\omega u$$

so that $T_\omega(u) = (\gamma(\omega_1)^*)^{-1}(u_\omega)$. It follows immediately that T_ω is a bounded linear operator for each ω in Ω and that the map $\omega \to T_\omega$ is continuous from Ω to $L(W, \Gamma)$ in the norm topology.

Finally, set $S_\omega = T_\omega^* \colon \Gamma \to W$. Then $\omega \to S_\omega$ is continuous from Ω to $L(\Gamma, W)$ in the norm-topology and

$$\gamma(\omega) S_\omega = \gamma(\omega) T_\omega^* = (T_\omega \gamma(\omega)^*)^* = I_\Gamma. \qquad \text{Q.E.D.}$$

LEMMA (1.7). *Let* $\{\Gamma, A(\omega), \gamma(\omega)\}$ *be a continuous family of abstract elliptic boundary value problems for* $\{H, W\}$. *For each* ω *in* Ω, *let* c_ω *be the least constant such that*

$$\| u \|_W \leqq c_\omega \{ \| u \|_H + \| A(\omega) u \|_H + \| \gamma(\omega) u \|_\Gamma \}$$

for all u *in* W.

Then for each ω_1 in Ω, there is a neighborhood N of ω_1 in Ω such that c_ω is bounded on N.

Proof of Lemma (1.7). By assumption, there exists a constant c such that

$$\| u \|_W \leqq c\{\| u \|_H + \| A(\omega_1)u \|_H + \| \gamma(\omega_1)u \|_\Gamma\}.$$

Hence

$$\begin{aligned}\| u \|_W \leqq c\{&\| u \|_H + \| A(\omega)u \|_H + \| (A(\omega_1) - A(\omega))u \|_H \\ &+ \| \gamma(\omega)u \|_\Gamma + \| (\gamma(\omega_1) - \gamma(\omega))u \|_\Gamma\} \\ \leqq c\{&\| u \|_H + \| A(\omega)u \|_H + \| \gamma(\omega)u \|_\Gamma\} \\ &+ \epsilon(\omega) \| u \|_W\end{aligned}$$

where $\epsilon(\omega) \to 0$ as $\omega \to \omega_1$. Choosing N so small that $\epsilon(\omega) < \frac{1}{2}$, we have $c_\omega \leqq 2c$ on N. Q.E.D.

LEMMA (1.8). *Under the hypotheses of Theorem 1, if*

$$\Omega_1 = \{\omega \mid \omega \in \Omega,\ (A'(\omega))_{\gamma'(\omega)} = (A(\omega)_{\gamma(\omega)})^*\},$$

then Ω_1 is closed in Ω.

Proof of Lemma (1.8). Let ω_1 be a point of the closure of Ω_1 in Ω. We wish to show that ω_1 lies in Ω_1.

Applying Lemma (1.7) and restricting our attention to a suitably small neighborhood of ω_1 in Ω, we may assume without loss of generality that there exists a fixed constant c such that for all u in W,

$$\| u \|_W \leqq c\{\| u \|_H + \| A(\omega)u \|_H + \| \gamma(\omega)u \|_\Gamma\}$$

and

$$\| u \|_W \leqq c\{\| u \|_H + \| A'(\omega)u \|_H + \| \gamma'(\omega)u \|_{\Gamma'}\}$$

for all ω in Ω.

Let u be an element of $D((A(\omega_1)_{\gamma(\omega_1)})^*)$ and let $w = (A(\omega_1)_{\gamma(\omega_1)})^*u$. Then for all v_1 in W with $\gamma(\omega_1)v_1 = 0$, we have

$$(u, A(\omega_1)v_1) = (w, v_1).$$

Let $\omega \in \Omega$, $v \in D((A(\omega))_{\gamma(\omega)})$. Set $v_1 = v - S_{\omega_1}\gamma(\omega_1)v$, where S_ω is the continuous family of right inverses for $\gamma(\omega)$ whose existence was proved in Lemma (1.6). Then $\gamma(\omega_1)(v_1) = 0$, and we have

$$(u, A(\omega_1)(v - S_{\omega_1}\gamma(\omega_1)v)) = (w, v) - (w, S_{\omega_1}\gamma(\omega_1)v). \tag{1.4}$$

We rewrite equation (1.4) as follows:

$$(1.5)\quad \begin{aligned}(u, A(\omega)v) &= (w, v) + (u, (A(\omega) - A(\omega_1))v)\\ &\quad + (u, A(\omega_1)S_{\omega_1}(\gamma(\omega_1) - \gamma(\omega))v) + (w, S_{\omega_1}(\gamma(\omega) - \gamma(\omega_1))v).\end{aligned}$$

Consider the sum

$$(1.6)\quad \begin{aligned}s(v) &= (u, (A(\omega) - A(\omega_1))v) + (w, S_{\omega_1}(\gamma(\omega) - \gamma(\omega_1))v)\\ &\quad + (u, A(\omega_1)S_{\omega_1}(\gamma(\omega) - \gamma(\omega))v).\end{aligned}$$

Using the continuity of $A(\omega)$ and $\gamma(\omega)$ in ω near ω_1, we have

$$|s(v)| \leqq \epsilon(\omega)\{\| u \|_H \| v \|_W + \| w \|_H \| v \|_W\}$$

where $\epsilon(\omega) \rightarrow 0$ as $\omega \rightarrow \omega_1$ independently of u, v, or w.

Since $\gamma(\omega)v = 0$, we have

$$\| v \|_W \leqq c\{\| v \|_H + \| A(\omega)v \|_H\}.$$

Thus

$$|s(v)| \leqq c\epsilon(\omega)\{\| u \|_H + \| w \|_H\}\{\| v \|_H + \| A(\omega)v \|_H\}$$

The linear functional $s(v)$ of $[v, A(\omega)v]$ on the graph of $(A(\omega))_{\gamma(\omega)}$ is therefore bounded with a bound $\leqq c\epsilon(\omega)\{\| u \|_H + \| w \|_H\}$. We may extend this functional to a linear functional on $H \oplus H$ with the same bound, i.e. there exist elements u_ω and w_ω of H with

$$(\| u_\omega \|^2{}_H + \| v_\omega \|^2{}_H)^{\frac{1}{2}} \leqq c\epsilon(\omega)\{\| u \|_H + \| w \|_H\}$$

such that for all v in W with $\gamma(\omega)v = 0$,

$$(1.7)\quad s(v) = (u_\omega A(\omega)v) + (w_\omega, v).$$

Substituting equation (1.7) in equations (1.5) and (1.6), we obtain

$$(u, A(\omega)v) = (w, v) + (u_\omega, A(\omega)v) + (w_\omega, v)$$

for all $v \in D((A(\omega))_{\gamma(\omega)})$. Thus,

$$(1.8)\quad (u - u_\omega, A(\omega)v) = (w + w_\omega, v).$$

If we consider only ω lying in Ω_1, equation (1.8) implies that

$$u - u_\omega \in D((A'(\omega))_{\gamma'(\omega)})$$

i.e. $u = u_\omega \in W$, $\gamma'(\omega)(u - u_\omega) = 0$, and

$$(1.9)\quad A'(\omega)(u - u_\omega) = (w + w_\omega)$$

We know that

$$\begin{aligned}\| u - u_\omega \|_W &\leqq c\{\| u - u_\omega \|_H + \| A'(\omega)(u - u_\omega)\|_H\} \\ &\leqq c\{\| u - u_\omega \|_H + \| w + w_\omega \|_H\} \\ &\leqq c\{\| u \|_H + \| w \|_H\}(1 + 2c\epsilon(\omega)).\end{aligned}$$

Thus the set of elements $\{u - u_\omega \mid \omega \in \Omega_1\}$ is bounded in W and the intersection of this set with the filter of neighborhoods of ω_1 is itself a filter base. By the weak compactness of closed balls in W, there exists an element u_0 of W which is an adherent point of the corresponding filter in the weak topology on W and hence in the strong topology on H. Since $\| u_\omega \| \to 0$ as $\omega \to \omega_1$, $u - u_\omega \to u$ in H. Hence $u = u_0$ and u lies in W. Furthermore by the weak compactness of bounded closed sets and the uniqueness of the adherent point u_0, $u - u_\omega \to u$ weakly in W as $\omega \to \omega_1$.

Since $\gamma'(\omega) \to \gamma'(\omega_1)$ in uniform norm as $\omega \to \omega_1$ while $u - u_\omega \to u$ boundedly and weakly in W, it follows that

$$\gamma'(\omega)(u - u_\omega) \to \gamma'(\omega_1)u$$

weakly in Γ. However, $\gamma'(\omega)(u - u_\omega) = 0$ for all ω in Ω_1 so that $\gamma'(\omega_1)u = 0$. Hence $D((A(\omega_1))^*_{\gamma(\omega_1)}) \subset D((A'(\omega_1))_{\gamma'(\omega_1)})$ and since $(A'(\omega_1))_{\gamma(\omega_1)}$ is a restriction of $(A(\omega_1))^*_{\gamma(\omega_1)}$, it follows that

$$(A(\omega_1)_{\gamma(\omega_1)})^* = (A'(\omega_1))_{\gamma'(\omega_1)}. \qquad \text{Q.E.D.}$$

Proof of Theorem 1. By Lemma (1.5) Ω_1 is open in Ω. By Lemma (1.8), Ω_1 is closed in Ω. Since Ω is connected and ω_0 lies in Ω_1, it follows that $\Omega_1 = \Omega$, i.e.

$$(A'(\omega))_{\gamma(\omega)} = (A(\omega)_{\gamma(\omega)})^*$$

for all ω in Ω. Q.E.D.

Section 2. Let X be a precompact manifold with boundary, Y its boundary, B a differentiable complex vector bundle over X. We may assume without loss of generality that X is differentiably imbedded in a compact manifold X_1 and that B is the restriction to X of a differentiable vector bundle B_1 over X_1. We suppose given a Hermitian structure on the fibres of B_1 and a differentiable measure μ on X_1. Following the discussion of Section 2 of [10], we may then form:

$\mathcal{E}(B)$ = the family of infinitely differentiable sections of B;
$\mathfrak{D}(B) = \{u \mid u \in \mathcal{E}(B)$, Supp$(u)$ is a compact subset of the interior of $X\}$;

$\mathfrak{D}'(B)$ = the family of distribution sections of B;
$L^2(B)$ = the Hilbert space formed from the measurable sections of of B, square-summable with respect to μ;
$W^{m,2}(B)$ = the space of sections of B with square-summable derivatives of order $\leqq m$.

(For the precise definitions, we refer to [10]. Note that the injection map of $W^{m,2}(B)$ into $L^2(B)$ is compact under our hypotheses, [25].)

If A is a linear differential operator of order m in B_1, then if we restrict A to B, we see that A defines a bounded linear mapping of $W^{m,2}(B)$ into $L^2(B)$.

Definition (2.1). *By an elliptic boundary value problem of order m on B, we mean a triple $\{A, \Gamma, \gamma\}$ consisting of:*

(1) *A continiuous linear mapping A of $W^{m,2}(B)$ into $L^2(B)$;*
(2) *A Hilbert space Γ;*
(3) *A continuous linear mapping γ of $W^{m,2}(B)$ onto Γ such that*

$$\{u \mid u \in W^{m,2}(B), \gamma u = 0\} \text{ is dense in } L^2(B);$$

and satisfying the basic condition: There exists a constant $c > 0$ such that for all u in $W^{m,2}(B)$,

$$\| u \|_{W^{m,2}(B)} \leqq c\{\| u \|_{L^2(B)} + \| Au \|_{L^2(B)} + \| \gamma u \|_{\Gamma}\}. \tag{2.1}$$

In the terminology of Section 1 of the present paper, an elliptic boundary value problem of order m on B is simply an abstract elliptic boundary value problem $\{\Gamma, A, \gamma\}$ for $\{L^2(B), W^{m,2}(B)\}$.

Definition (2.2). *The boundary value problem $\{A, \Gamma, \gamma\}$ is said to be proper if $\gamma u = 0$ for all u in $\mathfrak{D}(B)$. It is said to be local if $\Gamma \subset \mathfrak{D}'(B_2)$ where B_2 is a vector bundle over Y and γ is a local operator,* (*i.e. with* $\operatorname{Supp}(\gamma u) \subset \operatorname{Supp}(u)$ *for* $u \in \mathcal{E}(B)$).

The class of local elliptic boundary value problems where A is an elliptic differential operator has been studied intensively in recent years (cf. Chapter X of [16] for a brief development of basic results). This discussion has been extended also to semi-local problems, i.e. where A and γ are given by differential operators with singular integral operators as coefficients (see [2] for a summary of results).

We apply the results of Section 1 in a slight reformulation to obtain a general result which (as we shall show elsewhere in detail) is very useful in answering the following two types of questions for non-local elliptic boundary value problems:

(I) *If* (A,γ) *is the mapping of* $W^{m,2}(B)$ *into* $L^2(B)\oplus\Gamma$ *given by*

$$(A,\gamma)(u)=[Au,\gamma u],$$

then is (A,γ) *a Fredholm operator?*

(II) *If* A_γ *is the closed densely defined linear operator in* $L^2(B)$ *with*

$$D(A_\gamma)=\{u\mid u\in W^{m,2}(B),\gamma u=0\},$$
$$A_\gamma u=Au,u\in D(A_\gamma),$$

and if $\{A',\Gamma',\gamma'\}$ *is another elliptic boundary value problem which is formally adjoint to* $\{A,\Gamma,\gamma\}$ *in the sense that*

$$(Au,v)=(u,A'v)$$

for $u\in D(A_\gamma)$, $v\in D(A'_{\gamma'})$, *then is it true that*

$$A'_{\gamma'}=(A_\gamma)^*?$$

To present our general result, we follow Section 1 to define:

Definition (2.3). *Let* Ω *be a topological space and for each* ω *in* Ω, *let* $\{(A(\omega),\Gamma,\gamma(\omega)\}$ *be an elliptic boundary value problem of order* m *on* B *with a fixed Hilbert space* Γ. *Then this family of elliptic boundary value problems is said to be continuous if the mapping* f *of* Ω *into* $L(W^{m,2}(B), L^2(B)\oplus\Gamma)$ *given by* $f(\omega)(u)=[A(\omega)u,\gamma(\omega)u]$, *is continuous from* Ω *to the uniform topology on* $L(W^{m,2}(B),L^2(B)\oplus\Gamma)$.

THEOREM 2. *Let* Ω *be a connected topological space and let*

$$A(\omega),\Gamma,\gamma(\omega)\}\text{ and }\{A'(\omega),\Gamma',\gamma'(\omega)\}$$

be two continuous families of elliptic boundary value problems of order m *on* B *such that for each* ω *in* Ω, $\{A(\omega,\Gamma,\gamma(\omega)\}$ *is formally adjoint to* $\{A'(\omega),\Gamma',\gamma'(\omega)\}$. *Suppose further that there exists* ω_0 *in* Ω *such that* $(A'(\omega_0))_{\gamma'(\omega_0)}=((A(\omega_0))_{\gamma(\omega_0)})^*$.

Then

(a) *For all* ω *in* Ω,

$$(A'(\omega))_{\gamma(\omega)}=(A(\omega)_{\gamma(\omega)})^*.$$

(b) *$(A(\omega),\gamma(\omega))$ and $(A(\omega))_{\gamma(\omega)}$ are Fredholm operators for all ω in Ω. Their indexes are equal and independent of ω in Ω.*

Proof of Theorem 2. This is just the specialization in our present context of Theorem 1 of Section 1 with $H = L^2(B)$, $W = W^{m,2}(B)$.

Q. E. D.

INSTITUTE FOR ADVANCED STUDY
AND
UNIVERSITY OF CHICAGO.

REFERENCES.

[1] S. Agmon, A. Douglis and L. Nirenberg, "Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions, I," *Communications on Pure Applied Mathematics*, vol. 12 (1959), pp. 623-727.

[2] M. S. Agranovich and A. S. Dynin, "General boundary-value problems for elliptic systems in an *n*-dimensional domain," *Doklady Akademii Nauk SSSR*, vol. 146 (1962), pp. 511-514 (*Soviet Mathematics, Doklady*, vol. 3, No. 5, pp. 1323-1327).

[3] W. G. Bade and R. S. Freeman, "Closed extensions of the Laplace operator determined by a general class of boundary conditions," *Pacific Journal of Mathematics*, vol. 12 (1962), pp. 395-410.

[4] R. W. Beals, "A note on the adjoint of a perturbed operator," *Bulletin of the American Mathematical Society*, vol. 70 (1964), pp. 314-315.

[5] ———, *Non-local boundary value problems for elliptic differential operators*, Yale Ph. D. Thesis, January, 1964.

[6] F. Browder, "Estimates and existence theorems for elliptic boundary value problems," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 45 (1959), pp. 365-372.

[7] ———, "A priori estimates for solutions of elliptic boundary value problems, I, II, III," *Koninklijke Nederlandse Akademie van Wetensohappen, Proceedings, Series A*, vol. 22 (1960), pp. 145-159, 160-169; vol. 23 (1961), pp. 404-410.

[8] ———, "A continuity property for adjoints of closed operators in Banach spaces and its application to elliptic boundary value problems," *Duke Mathematical Journal*, vol. 28 (1961), pp. 157-182.

[9] ———, "Functional analysis and partial differential equations, II," *Mathematische Annalen*, vol. 145 (1962), pp. 81-226.

[10] ———, "Analyticity and partial differential equations, I," *American Journal of Mathematics*, vol. 84 (1962), pp. 666-710.

[11] ———, "Families of closed linear operators depending upon a parameter," (*to appear*).

[12] ———, "Functional analysis and partial differential equations, III" (*to appear*).

[13] J. W. Calkin, "Abstract symmetric boundary conditions," *Transactions of the American Mathematical Society*, vol. 45 (1939), pp. 360-442.

[14] R. S. Freeman, "On closed extensions of second order formally self-adjoint uniformly elliptic differential operators" (*to appear*).

[15] I. C. Gokhberg and M. G. Krein, "The basic propositions on defect numbers, root numbers, and indices of linear operators," *Uspehi Matematičeskih Nauk* (*N. S.*), vol. 12 (1957), pp. 43-118 (*American Mathematical Society Translation* (2), vol. 13 (1960), pp. 185-264).

[16] L. Hörmander, "On the theory of general partial differential operators," *Acta Mathematica*, vol. 94 (1955), pp. 161-248.

[17] ———, *Linear partial differential operators* (Die Grundlehren der Mathematischer Wissenschaften in Einzeldarstellungen, vol. 116), Berlin, 1963.

[18] J. L. Lions and E. Magenes, "Problemi ai limiti non omogenei, III, IV, V," *Annali di Pisa*, vol. 15 (1961), pp. 39-101, 311-326; vol. 16 (1962), pp. 1-44.

[19] Ya B. Lopatinski, "On a method of reducing boundary problems for a system of differential equations of elliptic type to regular equations," *Ukrainskiĭ Matematičeskiĭ Žurnal*, vol. 5 (1953), pp. 123-151.

[20] J. Peetre, *Théorèmes de régularité pour quelques classes d'opérateurs differentiels*, Thesis, Lund, 1959.

[21] ———, "Another approach to elliptic boundary problems," *Communications on Pure and Applied Mathematics*, vol. 14 (1961), pp. 711-731.

[22] R. S. Phillips, "Dissipative operators and hyperbolic systems of partial differential equations," *Transactions of the American Mathematical Society*, vol. 90 (1959), pp. 193-254.

[23] M. Schechter, "Integral inequalities for partial differential operators and functions satisfying general boundary conditions," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. 37-66.

[24] ———, "General boundary value problems for elliptic partial differential equations," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. 457-486.

[25] ———, "Various types of boundary conditions for elliptic equations," *Communications on Pure and Applied Mathematics*, vol. 13 (1960), pp. 407-425.

[26] R. T. Seeley, "Integro-differential operators on vector bundles, I" (*to appear*).

[27] M. I. Visik, "On general boundary value problems for elliptic partial differential equations," *Trudy Moskovskogo Matematičeskogo Obščestra*, vol. 1 (1952), pp. 187-246 (*American Mathematical Society Translation* (2), vol. 24 (1963), pp. 107-172).

ON THE STRUCTURE OF COMPACT COMPLEX ANALYTIC SURFACES, I.

By K. KODAIRA.

The present paper is a continuation of a series of our previous papers [8], [9]. The main results of this paper were announced in two short notes [10], [11].

We call a compact complex analytic surface free from singularities simply a *surface*. In our previous paper [8], we have shown that any surface with non-constant meromorphic functions is either an algebraic surface or an elliptic surface and, in [9], we have examined the structure of elliptic surfaces in detail.

In Section 1 of the present paper we establish certain relations between several numerical characters, e.g., the geometric genus, the irregularity, the Betti numbers, . . . of a surface (Theorem 3). In Section 2 we derive some sufficient conditions for surfaces to be algebraic or elliptic. In particular we prove that a surface with an even first Betti number is algebraic if its geometric genus vanishes (Theorem 10). Following A. Weil we define K3 surface to be a regular surface of which the first Chern class vanishes. In Section 5 we prove a theorem to the effect that every K3 surface is a deformation of a non-singular quartic surface in a projective 3-space (Theorem 13). This theorem has been conjectured earlier by A. Weil and independently by A. Andreotti (compare Grauert [5]). In Section 3 we examine those surfaces which have no (non-constant) meromorphic functions. By a theorem of Grauert [6], any surface is obtained from a surface containing no exceptional curve (of the first kind) by means of a finite number of quadric transformations. Hence, in order to study the structure of surfaces, it suffices to consider surfaces containing no exceptional curves. We prove that, if a surface containing no exceptional curve has no meromorphic function, then it is a complex torus, a K3 surface or a surface of which the first Betti number is 1 and the geometric genus is zero (Theorem 11). In Section 6 we show that a surface with a trivial canonical bundle is a complex torus, a K3 surface or an elliptic surface of the form $\boldsymbol{C}^2/G$, where $\boldsymbol{C}^2$ denotes the

Received May 19, 1964.

space of two complex variables and G is a properly discontinuous non-abelian group of affine transformations of $\boldsymbol{C}^2$ (Theorem 19). Finally, in Section 7, we introduce a classification of surfaces (Theorems 21 and 22). We prove that a surface is a deformation of an algebraic surface if and only if its first Betti number is even (Theorem 25).

The proof of our results is based on the Riemann-Roch-Hirzebruch theorem of which the complete form has been established recently by Atiyah and Singer [2].

1. Numerical characters. By a *surface* we shall mean a compact complex analytic surface free from singularities. Throughout this paper we denote by S a surface. In this section we shall establish certain relations between several *numerical characters,* e.g., the geometric genus, the irregularity, the Betti numbers, . . . of a surface.

We denote by b_ν the ν-th Betti number of S and by c_ν the ν-th Chern class of S. Moreover, for any sheaf Ξ over S, we denote by $H^\nu(S, \Xi)$ the ν-th cohomology group of S with coefficients in Ξ. We represent any cohomology class $c \in H^4(S, \boldsymbol{Z})$ by the value $c(S)$ of c on the fundamental cycle of the surface S *oriented in the natural way* with respect to its complex structure, where $\boldsymbol{Z}$ denotes the ring of (rational) integers, and we write c for $c(S)$. Thus we consider c_1^2 and c_2 as integers. We denote by $\mathcal{O}$ and Ω^r respectively the sheaves over S of germs of holomorphic functions and of holomorphic r-forms. We define

$$h^{r,\nu} = \dim H^\nu(S, \Omega^r), \qquad \Omega^0 = \mathcal{O}.$$

By the duality theorem $h^{0,2}$ is equal to $h^{2,0}$. We set

$$p_g = h^{0,2} = h^{2,0}, \qquad q = h^{0,1},$$

and we call p_g and q respectively the *geometric genus* and the *irregularity* of S.

Let $\boldsymbol{R}$ denote the field of real numbers. In view of de Rham's theorem each element of $H^2(S, \boldsymbol{R})$ is represented by a real d-closed 2-form defined on S. For any real d-closed 2-forms ξ and η on S we define

$$(\xi\eta) = \int_S \xi \wedge \eta.$$

The integral $(\xi\eta)$ represents the cup product of the cohomology classes of ξ and of η. Hence we may consider $(\xi\eta)$ as a non-singular symmetric bilinear form defined on the linear space $H^2(S, \boldsymbol{R})$. We define b^+ and b^- to be respectively the numbers of positive and of negative eigen-values of the symmetric bilinear form $(\xi\eta)$.

We write $b_\nu(S), h^{r,\nu}(S), p_g(S), \cdots$ for $b_\nu, h^{r,\nu}, p_g, \cdots$ when we want to make explicit that they are numerical characters of the surface S.

Let $\boldsymbol{\mathcal{O}}^*$ denote the multiplicative sheaf over S of germs of non-vanishing holomorphic functions. We have the exact sequence

$$0 \to \boldsymbol{Z} \to \boldsymbol{\mathcal{O}} \to \boldsymbol{\mathcal{O}}^* \to 0$$

and the corresponding exact cohomology sequence

$$(1) \quad \cdots \to H^1(S, \boldsymbol{\mathcal{O}}) \to H^1(S, \boldsymbol{\mathcal{O}}^*) \xrightarrow{\delta^*} H^2(S, \boldsymbol{Z}) \to H^2(S, \boldsymbol{\mathcal{O}}) \to \cdots.$$

Each element F of $H^1(S, \boldsymbol{\mathcal{O}}^*)$ represents a complex line bundle over S. The Chern class $c(F)$ is, by definition, the image $\delta^* F$ of F in $H^2(S, \boldsymbol{Z})$. For any complex line bundle F over S we denote by $\boldsymbol{\mathcal{O}}(F)$ the sheaf over S of germs of holomorphic sections of F. In the case of complex line bundles over surfaces, the Riemann-Roch-Hirzebruch theorem can be formulated as follows:

$$(2) \quad \sum_{\nu=0}^{2} (-1)^\nu \dim H^\nu(S, \boldsymbol{\mathcal{O}}(F)) = \tfrac{1}{2}(c^2 + c_1 c) + \tfrac{1}{12}(c_1{}^2 + c_2), \; c = c(F)$$

(see Atiyah and Singer [2], Hirzebruch [7]). Hence the following formula of Noether follows:

$$(3) \qquad 12(p_g - q + 1) = c_1{}^2 + c_2.$$

In what follows we shall employ the notation of our previous paper [8] for linear systems of divisors. Thus, letting F be a complex line bundle over S, we denote for any holomorphic section σ of F the *divisor* of σ by (σ), provided that σ does not vanish identically, and define $|F|$ to be the *complete linear system* consisting of all divisors (σ), $\sigma \in H^0(S, \boldsymbol{\mathcal{O}}(F))$, $\sigma \neq 0$. Moreover we denote by K the *canonical bundle* of S. We have

$$\dim |F| = \dim H^0(S, \boldsymbol{\mathcal{O}}(F)) - 1.$$

Hence, using the duality theorem, we obtain from (2) the Riemann-Roch inequality

$$(4) \quad \dim |F| + \dim |K - F| \geqq \tfrac{1}{2}(c^2 + c_1 c) + p_g - q - 1, \quad c = c(F).$$

We denote by (z_1, z_2) a local coordinate (not specified) on S. Since S is oriented in the natural way, the volume element $dz_1 \wedge dz_2 \wedge d\bar{z}_1 \wedge d\bar{z}_2$ is positive. Hence, for any differential form $\psi = \psi_{12} dz_1 \wedge dz_2$ on S which does does not vanish identically, we have

$$\int_S \psi \wedge \bar{\psi} > 0. \tag{5}$$

THEOREM 1. *Every holomorphic* 1*-form defined on a surface is d-closed.*

Proof. Let ϕ denote a holomorphic 1-form defined on S. We have

$$\int_S d\phi \wedge d\bar{\phi} = \int_S d(\phi \wedge d\bar{\phi}) = 0.$$

Comparing this with (5) we conclude that $d\phi$ vanishes identically.

THEOREM 2. *Let $\phi_1, \phi_2, \cdots, \phi_h$ be holomorphic* 1*-forms defined on S. If $\phi_1, \cdots, \phi_h$ are linearly independent, then the d-closed* 1*-forms $\phi_1, \cdots, \phi_h$, $\bar{\phi}_1, \cdots, \bar{\phi}_h$ are d-cohomologically independent.*

Proof. Assume that

$$\alpha_1\phi_1 + \cdots + \alpha_h\phi_h + \beta_1\bar{\phi}_1 + \cdots + \beta_h\bar{\phi}_h = d\mu, \tag{6}$$

where the coefficients $\alpha_1, \cdots, \alpha_h$, $\beta_1, \cdots, \beta_h$ are constants and μ denotes a continuously differentiable function defined on S. For any point z on S we find holomorphic functions f_j defined on a neighborhood of z in S such that $df_j = \phi_j$. We have

$$\mu = \alpha_1 f_1 + \cdots + \alpha_h f_h + \beta_1 \bar{f}_1 + \cdots + \beta_n \bar{f}_h + \text{constant}.$$

This shows that μ satisfies the mean value theorem with respect to any sufficiently small "sphere" in S. Hence μ is reduced to a constant and consequently

$$\alpha_1\phi_1 + \cdots + \alpha_h\phi_h = 0, \qquad \bar{\beta}_1\phi_1 + \cdots + \bar{\beta}_n\phi_h = 0. \tag{7}$$

Thus we see that the d-cohomological relation (6) is equivalent to the linear relations (7), q.e.d.

Let $\boldsymbol{C}$ denote the field of complex numbers. We have the exact sequence

$$0 \to \boldsymbol{C} \to \mathcal{O} \xrightarrow{d} d\mathcal{O} \to 0$$

and the corresponding exact cohomology sequence

$$0 \to H^0(S, d\mathcal{O}) \to H^1(S, \boldsymbol{C}) \to H^1(S, \mathcal{O}) \to \cdots. \tag{8}$$

Hence we get the inequality

$$q \geqq b_1 - \dim H^0(S, d\mathcal{O}).$$

On the other hand we infer from Theorems 1 and 2 that

$$h^{1,0} = \dim H^0(S, d\mathcal{O}) \leqq \tfrac{1}{2}b_1. \tag{9}$$

Consequently we obtain

$$q \geqq b_1 - h^{1,0} \geqq \tfrac{1}{2}b_1. \tag{10}$$

Moreover we have the *inequality*

$$b^+ \geqq 2p_g. \tag{11}$$

In fact, in view of (5), we can choose p_g holomorphic 2-forms ψ_j, $j = 1, 2, \cdots, p_g$, such that

$$\int \psi_j \wedge \bar{\psi}_k = \delta_{jk}.$$

Hence, setting

$$\eta_{2j-1} = \bar{\psi}_j + \psi_j, \qquad \eta_{2j} = i\bar{\psi}_j - i\psi_j, \qquad i = \sqrt{-1},$$

we obtain

$$(\eta_j\eta_k) = 2\delta_{jk}, \qquad \text{for } j, k = 1, 2, \cdots, 2p_g.$$

This proves the inequality (11).

By the Hirzebruch index theorem (see Hirzebruch [7]) we have

$$3b^+ - 3b^- = c_1^2 - 2c_2. \tag{12}$$

Combining this with (3) and using the identity

$$c_2 = 2 - 2b_1 + b^+ + b^-,$$

we obtain

$$2q - b_1 + b^+ - 2p_g = 1,$$

while, by (10) and (11), the integers $2q - b_1$ and $b^+ - 2p_g$ are both non-negative. Hence we conclude that $2q$ is equal to either b_1 or $b_1 + 1$ and that b^+ is equal to $2p_g + 1$ or $2p_g$ according as $2q = b_1$ or $= b_1 + 1$. Thus we obtain the following

THEOREM 3. *If b_1 is even, then $2q = b_1$, $2p_g = b^+ - 1$ and $h^{1,0} = q$. If b_1 is odd, then $2q = b_1 + 1$, $2p_g = b^+$ and $h^{1,0} = q - 1$.*

COROLLARY. *The numerical characters $p_g(S)$, $q(S)$ and $h^{1,0}(S)$ are homological invariants of the surface S.*

We obtain from (3) and Theorem 3 the formula

$$c_1^2 + 8q + b^- = \begin{cases} 10p_g + 9 & \text{if } b_1 \text{ is even,} \\ 10p_g + 8, & \text{if } b_1 \text{ is odd.} \end{cases} \tag{13}$$

Since, by Theorem 3, $b_1 = h^{1,0} + q$, we infer from (8) and (9) the *exactness of the sequence*

$$0 \to H^0(S, d\mathbf{O}) \to H^1(S, \mathbf{C}) \to H^1(S, \mathbf{O}) \to 0. \tag{14}$$

Now we consider the case in which the first Betti number b_1 of S is *even*. By the above results $b_1 = 2q$, $h^{1,0} = q$ and there exist q linearly independent d-closed holomorphic 1-forms $\phi_1, \phi_2, \cdots, \phi_q$ defined on S. Let $\{\gamma_1, \gamma_2, \cdots, \gamma_{2q}\}$ be a Betti base of 1-cycles on S and let

$$\omega_{\nu j} = \int_{\gamma_j} \phi_\nu.$$

Moreover let $\mathbf{C}^q$ denote the space of q complex variables. By Theorem 2 the d-closed 1-forms $\phi_1, \cdots, \phi_q, \bar{\phi}_1, \cdots, \bar{\phi}_q$ form a base of the d-cohomology group of 1-forms on S. Hence the vectors

$$\omega_j = (\omega_{1j}, \cdots, \omega_{\nu j}, \cdots, \omega_{qj}), \qquad j = 1, 2, \cdots, 2q,$$

are linearly independent over $\mathbf{R}$ and generate a discontinuous subgroup $\mathfrak{D}$ of the vector group $\mathbf{C}^q$. We call the compact complex analytic group variety $\mathfrak{A} = \mathbf{C}^q/\mathfrak{D}$ the *Albanese variety* attached to the surface S and define a *holomorphic map* Φ of S into $\mathfrak{A}$ by the formula

$$z \to \Phi(z) = \left(\int^z \phi_1, \cdots, \int^z \phi_q \right) \quad (\mathrm{mod}\ \mathfrak{D}). \tag{15}$$

For any 2-cycles Γ and Δ on a surface we denote by $(\Gamma\Delta)$ the *intersection multiplicity* of Γ and Δ and write (Γ^2) for $(\Gamma\Gamma)$. An *exceptional curve* (of the first kind) on a surface is, by definition, a non-singular rational curve C with $(C^2) = -1$. For any point w on a surface W we denote by Q_w the *quadric transformation of the center* w. The *quadric transform* $Q_w(W)$ of W is a surface and the *total transform* $Q_w(w)$ of w is an exceptional curve on $Q_w(W)$. It has been shown by H. Grauert that, conversely, *if C is an exceptional curve on a surface S, then there exist a surface W and a point $w \in W$ such that $S = Q_w(W)$ and $C = Q_w(w)$* (see Grauert [6]. For an elementary proof of this result, see Kodaira [9], II, Appendix). Consequently *any surface is obtained from a surface containing no exceptional curve by means of a finite number of quadric transformations.*

The numerical characters b_1 and b^+ of a surface are invariant under quadric transformations. Hence, by Theorem 3, *the geometric genus p_g and the irregularity q of a surface are invariant under quadric transformations.*

2. Algebraic and elliptic surfaces.

By an *elliptic surface* we shall mean an analytic fibre space of elliptic curves over a non-singular algebraic curve, i. e., a surface S together with a holomorphic map Ψ of S onto a non-singular algebraic curve Δ such that the inverse image $\Psi^{-1}(u)$ of any general point $u \in \Delta$ is an elliptic curve. We call Δ the *base curve* of the elliptic surface S.

THEOREM 4. *If there exist on S two algebraically independent meromorphic functions, then S is an algebraic surface. If there exists on S one and only one algebraically independent meromorphic function, then S is an elliptic surface* (see Kodaira [8], Chow and Kodaira [3]).

If there exists on S a complex line bundle F which has at least two linearly independent global holomorphic sections σ_1 and σ_2, then the quotient σ_1/σ_2 is a non-constant meromorphic function on S and, consequently, S is either algebraic or elliptic. Thus we obtain the following two theorems:

THEOREM 5. *If there exists on S a complex line bundle F such that* $\dim |F| \geqq 1$, *then S is either an algebraic surface or an elliptic surface.*

THEOREM 6. *If $p_g(S) \geqq 2$, then S is either an algebraic surface or an elliptic surface.*

THEOREM 7. *If $h^{1,0}(S) \geqq 3$, then S is either an algebraic surface or an elliptic surface.*

Proof. By hypothesis there exist at least three linearly independent holomorphic 1-forms ϕ_1, ϕ_2, ϕ_3 defined on S. If the holomorphic 2-forms $\phi_2 \wedge \phi_1$ and $\phi_3 \wedge \phi_1$ are linearly independent, then, by Theorem 6, S is either algebraic or elliptic. On the other hand, if there is a non-trivial linear relation

$$\alpha\phi_2 \wedge \phi_1 + \beta\phi_3 \wedge \phi_1 = 0,$$

where α and β are constants, then we have

$$\alpha\phi_2 + \beta\phi_3 = f\phi_1.$$

Thus we obtain a non-constant meromorphic function f defined on S. Hence S is algebraic or elliptic.

THEOREM 8. *If there exists on S a complex line bundle F with $c(F)^2 > 0$, then S is an algebraic surface.*

Proof. Since by hypothesis, $c(F)^2$ is a positive integer, for any sufficiently large integer m, we obtain from (4) the inequality

$$\dim|mF| + \dim|K - mF| \geqq \tfrac{1}{4}c(F)^2m^2.$$

This shows that either $|mF|$ or $|K - mF|$ contains positive divisors. Let D_m denote one of these divisors. Then, for all sufficiently large integers m and n, we obtain

$$(16) \qquad \dim|nD_m| \geqq \tfrac{1}{4}c(F)^2n^2m^2.$$

Combined with Theorem 5 this proves that S is either algebraic or elliptic.

Suppose that S is a non-algebraic elliptic surface. Then all meromorphic functions on S are induced from meromorphic functions on the base curve Δ of S (see Kodaira [8], §4). It follows that, for any divisor D on S and for any positive integer n, the inequality

$$\dim|nD| \leqq an$$

holds, where a is a constant. This contradicts (16), q. e. d.

Since $c_1 = c(-K)$, we obtain from Theorem 8 the following

THEOREM 9. *If $c_1{}^2(S) > 0$, then S is an algebraic surface.*

THEOREM 10. *If $b_1(S)$ is even and if $p_g(S)$ vanishes, then S is an algebraic surface.*

Proof. Since b_1 is even, the vanishing of p_g implies, by Theorem 3, that $b^+ = 1$. There exists therefore a cohomology class $c \in H^2(S, \mathbf{Z})$ with $c^2 > 0$. Since $H^2(S, \mathcal{O})$ vanishes, we infer from the exact sequence (1) the existence of a complex line bundle F over S with $c(F) = c$. Hence, using Theorem 8, we conclude that S is an algebraic surface.

3. Surfaces with no meromorphic functions.

A surface is said to be *regular* if and only if its irregularity vanishes.

Definition. By a K3 *surface we shall mean a regular surface of which the first Chern class vanishes.*

In case S is regular, the exact sequence (1) is reduced to

$$(17) \qquad 0 \to H^1(S, \mathcal{O}^*) \to H^2(S, \mathbf{Z}) \to H^2(S, \mathcal{O}) \to \cdots$$

and, consequently, each complex line bundle F over S is determined uniquely by its Chern class $c(F)$. In particular, the canonical bundle of S is trivial if and only if the first Chern class c_1 of S vanishes. Hence *a regular surface is a* K3 *surface if and only if its canonical bundle is trivial.*

THEOREM 11. *Assume that there exists on S no meromorphic function except constants and that S contains no exceptional curve. Then the irregularity q of S is not greater than* 2. *If q is equal to* 2, *then S is a complex torus. If q is equal to* 1, *then the first Betti number b_1 of S is equal to* 1 *and the geometric genus p_g of S vanishes. If q is equal to* 0, *then S is a* K3 *surface.*

Proof. In view of Theorem 6, the geometric genus p_g of S is equal to either 1 or 0. In case $p_g = 1$ we denote by ψ a holomorphic 2-form on S which does not vanish identically. The divisor (ψ) of ψ is a canonical divisor. We then denote by K the *canonical divisor* (ψ) instead of the canonical bundle. Moreover, if $K \neq 0$, we write

$$K = \sum_\nu k_\nu C_\nu,$$

where the C_ν denote the irreducible curves on S and the coefficients k_ν are positive integers.

LEMMA 1. *If $p_g = 1$, then $c_1{}^2 = 0$. If, moreover, $K \neq 0$, then the virtual genus $\pi(C_\nu)$ of each irreducible component C_ν of K is equal to either* 1 *or* 0 *and*

$$(C_\nu{}^2) = 2\pi(C_\nu) - 2. \tag{18}$$

To prove this lemma it suffices to consider the case in which $K \neq 0$. We have the equalities

$$(C_\nu{}^2) + (KC_\nu) = 2\pi(C_\nu) - 2, \tag{19}$$

$$(KC_\nu) = k_\nu(C_\nu{}^2) + \sum_{\lambda \neq \nu} k_\lambda(C_\lambda C_\nu). \tag{20}$$

Hence we infer that $(KC_\nu) \geqq 0$. In fact, if $(KC_\nu) < 0$, then, by (20), $(C_\nu{}^2) < 0$ and therefore, by (19), $(C_\nu{}^2) = -1$, $\pi(C_\nu) = 0$. This contradicts the assumption that S contains no exceptional curve. Now, in view of Theorem 9, we have

$$\sum k_\nu(KC_\nu) = (K^2) = c_1{}^2 \leqq 0.$$

Hence we obtain

$$c_1{}^2 = 0, \qquad (KC_\nu) = 0,$$

and therefore, by (19) and (20),

$$2\pi(C_\nu) - 2 = (C_\nu{}^2) \leqq 0.$$

Consequently $\pi(C_\nu)$ is equal to 1 or 0.

A) *The case in which b_1 is even.* It follows from Theorems 3, 6, 7 and 10 that $b_1 = 2q$, $q = h^{1,0} \leqq 2$ and $p_g = 1$.

i) *$h^{1,0}$ is equal to either* 2 *or* 0. In fact, if $h^{1,0}$ were equal to 1, then the Albanese variety $\mathfrak{A}$ attached to S would be an elliptic curve and non-constant meromorphic functions on $\mathfrak{A}$ would induce non-constant meromorphic functions on S. This contradicts our hypothesis.

ii) In case $q = h^{1,0} = 2$, there exist on S two linearly independent holomorphic 1-forms ϕ_1 and ϕ_2. Since there exists on S no meromorphic function except constants, the linear independence of ϕ_1 and ϕ_2 implies that the exterior product $\phi_1 \wedge \phi_2$ does not vanish identically. Hence the holomorphic map Φ defined by (15) maps S onto the Albanese variety $\mathfrak{A}$. We infer in the same manner as in the proof of Theorem 5.3 in our previous paper [8] that Φ is a biholomorphic map of S onto $\mathfrak{A}$. Consequently *S is a complex torus.*

iii) Since $p_g = 1$, it follows from the above Lemma 1 that $c_1^2 = 0$, while, in view of Theorem 5, $\dim |2K|$ is non-positive. Hence, in case $q = 0$, we obtain from (4) the inequality

$$\dim \ |-K| \geqq -\dim |2K| \geqq 0.$$

This proves that the effective divisor K vanishes and, consequently, *S is a* K3 *surface.*

B) *The case in which b_1 is odd.* It follows from Theorems 3, 6, 7 and 9 that $b_1 = 2q - 1$, $q = h^{1,0} + 1$, $b^+ = 2p_g$, $h^{1,0} \leqq 2$, $p_g \leqq 1$ and $c_1^2 \leqq 0$.

i) Suppose that $h^{1,0} = 2$. Then there exist on S two linearly independent holomorphic 1-forms ϕ_1 and ϕ_2, $q = 3$ and $b_1 = 5$. Since there exists on S no meromorphic function except constants, $\phi_1 \wedge \phi_2$ does not vanish identically. Hence $p_g = 1$ and therefore, by Lemma 1, $c_1^2 = 0$. The formula (13) then proves that $b^- = -6$. This is a contradiction.

ii) Suppose that $h^{1,0} = 1$. Then there exists on S a d-closed holomorphic 1-form ϕ which does not vanish identically, $q = 2$ and $b_1 = 3$. If $p_g = 1$, then we choose a holomorphic 2-form ψ on S such that

$$\int_S \psi \wedge \bar{\psi} = 1$$

and let

$$\eta_1 = \psi + \bar{\psi}, \qquad \eta_2 = i\bar{\psi} - i\psi, \qquad i = \sqrt{-1}.$$

By a (r,s)*-form* we mean a differential form of type (r,s). For any 1-form τ we denote by τ_0 and by τ_1 respectively the component of τ of type $(1,0)$ and of type $(0,1)$. Now we prove *the existence of a d-closed* 1*-form* σ *on* S *such that* $\phi,\bar{\phi}$ *and* σ *generate the d-cohomology group of* 1*-forms on* S *and such that*

(21) $$d\sigma_0 = \phi \wedge \bar{\phi}.$$

Let

$$\xi = i\phi \wedge \bar{\phi}.$$

Obviously ξ is a real d-closed 2-form on S. Since

$$\int_S \phi \wedge \bar{\phi} \wedge \phi \wedge \bar{\phi} = \int_S \phi \wedge \bar{\phi} \wedge \psi = 0,$$

we have

$$(\xi\xi) = (\xi\eta_1) = (\xi\eta_2) = 0,$$

while $b^+ = 2p_g$, $p_g = 1$ or 0 and, in case $p_g = 1$, $(\eta_j\eta_k) = 2\delta_{jk}$ for $j,k = 1,2$. Hence we infer the existence of a real 1-form ρ on S such that

(22) $$d\rho = \xi = i\phi \wedge \bar{\phi}.$$

b_1 is equal to 3 and, by Theorem 2, ϕ and $\bar{\phi}$ are d-cohomologically independent. Hence we can choose a C^∞-differentiable real d-closed 1-form τ on S such that ϕ, $\bar{\phi}$ and τ generate the d-cohomology group of 1-forms on S. With the aid of the Dolbeault isomorphism (see Dolbeault [4]) we infer from the exact sequence (14) that $\bar{\phi}$ and τ_1 generate the $\bar{\partial}$-cohomology group of $(0,1)$-forms on S. It follows from (22) that $\bar{\partial}\rho_1 = 0$. Hence we find constants a, b and a C^∞-differentiable function λ on S such that

$$i\rho_1 = a\bar{\phi} + b\tau_1 + \bar{\partial}\lambda.$$

We define

$$\sigma = \sigma_0 + \bar{\sigma}_0, \qquad \sigma_0 = -i\rho + a\bar{\phi} + b\tau + d\lambda.$$

Clearly σ_0 is a $(1,0)$-form and

$$d\sigma_0 = -i\,d\rho = \phi \wedge \bar{\phi}.$$

We have

$$\sigma = a\bar{\phi} + \bar{a}\phi + (b + \bar{b})\tau + d\mu, \qquad \mu = \lambda + \bar{\lambda}.$$

Thus σ is a real d-closed 1-form on S. To prove that ϕ, $\bar{\phi}$ and σ are d-cohomologically independent it suffices to verify that $b + \bar{b} \neq 0$. Suppose that $b + \bar{b} = 0$. Then we obtain

$$\sigma_0 = \bar{a}\phi + \partial\mu$$

and consequently

$$\phi \wedge \bar{\phi} - d\sigma_0 = \bar{\partial}\partial\mu.$$

The function μ attains its minimum at a point o on S. The integral $w(z) = \int_o^z \phi$ gives a holomorphic function $w = w(z)$ defined on a neighborhood of o in S. We have

$$\partial\bar{\partial}(|w|^2 + \mu) = 0.$$

Hence $|w|^2 + \mu$ is the real part of a holomorphic function and satisfies the mean value theorem with respect to any sufficiently small sphere of the center o, while

$$|w(z)|^2 \geqq |w(o)|^2 = 0, \qquad \mu(z) \geqq \mu(o).$$

Consequently w vanishes identically. This contradicts the assumption that $\phi \neq 0$. Thus we see that ϕ, $\bar{\phi}$ and σ generate the d-cohomology group of 1-forms on S.

Let W denote the universal covering of the surface S. The 1-forms ϕ and σ induce 1-forms on W. We use the same symbols ϕ and σ, respectively, to denote those induced 1-forms. By integrating ϕ we get a holomorphic function w_1 on W such that $dw_1 = \phi$. It follows from (21) that

$$d(\sigma_0 + \bar{w}_1\phi) = 0.$$

Hence $\sigma_0 + \bar{w}_1\phi$ is a d-closed holomorphic 1-form on W. Integrating this 1-form we get a holomorphic function w_2 on W such that $dw_2 = \sigma_0 + \bar{w}_1\phi$. Thus we obtain a pair of holomorphic functions w_1, w_2 on W satisfying the simultaneous differential equations

$$\left\{\begin{aligned} dw_1 &= \phi, \\ dw_2 &= \sigma_0 + \bar{w}_1\phi. \end{aligned}\right. \tag{23}$$

We denote by G the group of covering transformations of W over S. For any function $f = f(w)$ on W, where w denotes a "variable" point on W, and for any covering transformation $g \in G$, we define

$$(gf)(w) = f(gw).$$

We infer from (23) that

$$\begin{aligned} d(gw_1 - w_1) &= 0, \\ d(gw_2 - w_2) &= (g\bar{w}_1 - \bar{w}_1)\,dw_1. \end{aligned}$$

Hence we obtain the formula

$$\left\{\begin{aligned} gw_1 &= w_1 + \alpha, \\ gw_2 &= w_2 + \bar{\alpha}w_1 + \beta, \end{aligned}\right. \tag{24}$$

where α and β are constants depending on g.

The exterior product $dw_1 \wedge dw_2$ *does not vanish identically.* To prove this we suppose that $dw_1 \wedge dw_2$ vanishes identically. Then we have

$$dw_2 = f\, dw_1,$$

where f is a meromorphic function on W. Using (24) we get

$$gf = f + \bar{\alpha}. \tag{25}$$

It follows that df is a meromorphic 1-form *defined on* S. Since $df \wedge dw_1 = ddw_2 = 0$, we have

$$df = k\, dw_1,$$

where k is a meromorphic function on S. Hence k is reduced to a constant and therefore

$$f = kw_1 + \text{constant.}$$

Combined with (24) and (25), this proves that $kw_1 - \bar{w}_1$ is a function defined on S. This contradicts the d-cohomological independence of $\phi = dw_1$ and $\bar{\phi} = d\bar{w}_1$. Thus we see that

$$dw_1 \wedge dw_2 = \phi \wedge \sigma_0$$

is a holomorphic 2-form on S which does not vanish identically. It follows that $p_g = 1$.

Next, we shall prove that $dw_1 \wedge dw_2$ *does not vanish at each point on* S. Suppose that the divisor K of $dw_1 \wedge dw_2$ is not equal to zero and write

$$K = \sum k_\nu C_\nu.$$

Then, by the above lemma, the virtual genus $\pi(C_\nu)$ of each irreducible curve C_ν is equal to 1 or 0. Hence C_ν is either a non-singular elliptic curve or a rational curve possibly with a singular point.

The function w_1 *is constant on* C_ν. To prove this we suppose that w_1 is not constant on C_ν. Then dw_1 induces on C_ν a holomorphic 1-form which does not vanish identically. Hence C_ν must be a non-singular elliptic curve. We denote by U_ν a connected component of the inverse image of C_ν in W and consider w_1 and w_2 as functions on U_ν by restricting their domains. Then we have

$$dw_2 = f\, dw_1,$$

where f is a holomorphic function on U_ν. Using (24) we get

$$gf = f + \bar{\alpha}.$$

Hence df is a holomorphic 1-form on C_ν and therefore

$$df = k\, dw_1,$$

where k is a constant. It follows that $kw_1 - \bar{w}_1$ is a function defined on C_ν. This contradicts the d-cohomological independence of dw_1 and $d\bar{w}_1$ on C_ν.

The function w_2 *is constant on* C_ν. To prove this we suppose that w_2 is not constant on C_ν and derive a contradiction. Since w_1 is constant on C_ν we infer from (23) that dw_2 induces on C_ν a holomorphic 1-form which does not vanish identically. Hence C_ν is a non-singular elliptic curve. We take a 1-cycle γ on C_ν such that

$$\int_\gamma \sigma = \int_\gamma d(w_2 + \bar{w}_2) \neq 0$$

and choose a Betti base $\{\gamma_1, \gamma_2, \gamma_3\}$ of 1-cycle on S such that γ is homologous to $m\gamma_3$ with respect to rational coefficients, where m is an integer $\neq 0$. Then we have

$$(26) \qquad \int_{\gamma_3} \phi = m^{-1} \int_\gamma dw_1 = 0.$$

Let

$$\omega_j = \int_{\gamma_j} \phi, \qquad \text{for } j = 1, 2.$$

Since ϕ and $\bar{\phi}$ are d-cohomologically independent on S, the periods ω_1 and ω_2 are linearly independent over $\boldsymbol{R}$ and generate a discontinuous subgroup $\mathfrak{D}$ of $\boldsymbol{C}$. Moreover, in view (26), $z \to w_1(z) = \int^z \phi$ defines a holomorphic map of S onto the elliptic curve $\boldsymbol{C}/\mathfrak{D}$. This contradicts the assumption that there exists on S no meromorphic function except constants.

Consider a connected component C of the curve $\sum C_\nu$ and let N be a sufficiently small neighborhood of C in S. Since w_1 and w_2 are constant on C, we infer from (24) that w_1 *and* w_2 *are single-valued functions on* N. Writing $w_1 = w_1(z)$, $w_2 = w_2(z)$, we obtain therefore a holomorphic map

$$h : z \to h(z) = (w_1(z), w_2(z))$$

of N into the space $\boldsymbol{C}^2$ of the complex variables w_1 and w_2. h maps C onto a point $\mathfrak{p} \in \boldsymbol{C}^2$ and the "jacobian" $dw_1(z) \wedge dw_2(z)$ of h does not vanish at each point z of $N - C$. Hence we infer by an elementary consideration that $h(N)$ contains a neighborhood of $\mathfrak{p}$ in $\boldsymbol{C}^2$ and that $h^{-1}(\mathfrak{p}) = C$. Choose a spherical neighborhood $M_\epsilon \subset h(N)$ of $\mathfrak{p}$ of sufficiently small radius ϵ in

$\boldsymbol{C}^2$ and let $N_\epsilon = h^{-1}(M_\epsilon)$. Then h maps $N_\epsilon - C$ biholomorphically onto $M_\epsilon - \mathfrak{p}$, since $N_\epsilon - C$ forms an unramified covering manifold of the simply connected domain $M_\epsilon - \mathfrak{p}$ with respect to the covering map h. Hence h^{-1} is a *local modification*. Consequently h^{-1} is composed of quadric transformations and therefore one of the irreducible components of $C = h^{-1}(\mathfrak{p})$ is an exceptional curve. This contradicts our hypothesis. Thus we conclude that $dw_1 \wedge dw_2$ does not vanish at each point of S.

We fix a point o on S, choose three closed continuous curves $\gamma_1, \gamma_2, \gamma_3$ starting and ending at o which form a Betti base of 1-cycles on S and let

$$\alpha_j = \int_{\gamma_j} \phi, \qquad s_j = \int_{\gamma_j} \sigma, \qquad j = 1, 2, 3. \tag{27}$$

Then the vectors (α_j, s_j), $j = 1, 2, 3$, are linearly independent over $\boldsymbol{R}$ and generate a discontinuous subgroup $\mathfrak{D}$ of $\boldsymbol{C} \times \boldsymbol{R}$. Obviously the quotient space

$$T = \boldsymbol{C} \times \boldsymbol{R}/\mathfrak{D}$$

is a torus of topological dimension 3. Setting

$$u(z) = w_2(z) + \overline{w_2(z)} - |w_1(z)|^2,$$

we get

$$du(z) = \sigma.$$

Moreover $dw_1 \wedge dw_2$ does not vanish at each point of S. Hence the formula

$$z \to P(z) = (w_1(z), u(z)) \quad (\operatorname{mod} \mathfrak{D})$$

defines a differentiable map P of S *onto* T and S is a differentiable fibre bundle over T. Moreover each fibre $P^{-1}(t)$, $t \in T$, of S is connected, since the projections $P(\gamma_j)$ of γ_j, $j = 1, 2, 3$, generate the fundamental group of T. Consequently *S is a differentiable fibre bundle of circles over T.* We denote by γ_0 the fibre $P^{-1}(P(o))$ which passes through o.

Each closed continuous curve γ_j determines an element g_j of the fundamental group $\pi_1(S) = \pi_1(S, o)$ of S. For the sake of simplicity we identify $\pi_1(S)$ with the covering transformation group G in a natural manner and consider g_j as a covering transformation of W over S. Then, in view of (24) and (27), we have

$$\begin{cases} g_j w_1 = w_1 + \alpha_j, \\ g_j w_2 = w_2 + \bar{\alpha}_j w_1 + \beta_j, \end{cases} \qquad \text{for } j = 0, 1, 2, 3, \tag{28}$$

where α_0 is equal to zero.

The projection P induces a homomorphism of $\pi_1(S)$ onto the fundamental group $\pi_1(T)$ of T whose kernel is the cyclic subgroup of $\pi_1(S)$ generated by g_0 and $\pi_1(T)$ is abelian. Hence we have

$$g_j g_k g_j^{-1} g_k^{-1} = g_0^{n_{jk}},$$

where the exponent n_{jk} is an integer. Combining this with (28) we obtain

$$\bar{\alpha}_j \alpha_k - \bar{\alpha}_k \alpha_j = n_{jk} \beta_0. \tag{29}$$

Since the vectors (α_j, s_j), $j = 1, 2, 3$, are linearly independent over $\boldsymbol{R}$, we infer from (29) that

$$n_{23} s_1 + n_{31} s_2 + n_{12} s_3 \neq 0,$$
$$n_{23} \alpha_1 + n_{31} \alpha_2 + n_{12} \alpha_3 = 0.$$

Hence we conclude that the periods α_1, α_2, α_3 generate a discontinuous subgroup D of $\boldsymbol{C}$ such that $\boldsymbol{C}/D$ is compact and that $z \to w_1(z) \pmod{D}$ defines a holomorphic map of S onto the elliptic curve $\boldsymbol{C}/D$. This contradicts the assumption that there exists on S no non-constant meromorphic function.

iii) Thus we see that $h^{1,0} = 0$ and consequently, $q = b_1 = 1$. For our purpose it suffices therefore to show that $p_g = 0$. The *Picard variety* $\mathfrak{P}$ *attached to* S is, by definition, the subgroup of $H^1(S, \mathcal{O}^*)$ consisting of all complex line bundles F over S with $c(F) = 0$. Since $q = b_1 = 1$, the exact sequence (1) is reduced to

$$0 \to \boldsymbol{Z} \to \boldsymbol{C} \to H^1(S, \mathcal{O}^*) \to H^2(S, \boldsymbol{Z}) \to \cdots.$$

Hence the Picard variety $\mathfrak{P}$ is isomorphic to the Lie group $\boldsymbol{C}/\boldsymbol{Z}$.

Suppose that $p_g = 1$. Then, for *each* complex line bundle $F \in \mathfrak{P}$, we obtain from (4) the inequality

$$\dim |F| + \dim |K - F| \geqq -1.$$

Hence we infer that either $|F|$ or $|K - F|$ is non-empty. On the other hand the non-existence of non-constant meromorphic function on S implies that the number of irreducible curves of S is finite (see Kodaira [8], Theorem 5.1). It follows that there exist on S at most countably many non-empty complete linear systems. This contradicts the above results.

4. Multiple singular fibres of elliptic surfaces. In this section we shall examine the structure of neighborhoods of multiple singular fibres of elliptic surfaces. All possible types of multiple singular fibres can be listed as types ${}_mI_b$, $m = 2, 3, 4, \cdots$, $b = 0, 1, 2, \cdots$ (see Kodaira [9], Theorem

6.2). In what follows we write I_b in place of ${}_1I_b$ and, for the sake of simplicity, we mean by a fibre of type I_0 a *regular* fibre.

Let S be an anayltic fibre space of elliptic curves over a non-singular algebraic curve Δ of which the fibre C_a over a point $a \in \Delta$ is of type ${}_mI_b$, $m \geqq 2$. Let τ be a local uniformization variable of the center a on Δ and let E_a denote the disk: $|\tau| < \epsilon^m$, where ϵ is a small positive number. Take a disk D: $|\sigma| < \epsilon$, define F to be the analytic fibre space of elliptic curves over D induced from $S \mid E_a$ by means of the holomorphic map: $\sigma \rightarrow \tau = \sigma^m$ of D onto E_a, where $S \mid E_a$ denotes the restriction of S to E_a, and let F_σ denote the fibre of F over σ (see Kodaira [9], § 6). The fibre F_σ is regular for $\sigma \neq 0$ and the fibre F_0 is of type I_{mb}. Moreover F is a fintie unramified covering manifold of $S \mid E_a$ whose covering transformation group is a cyclic group $\mathfrak{G}$ of order m generated by an automorphism g of F such that

$$gF_\sigma = F_{\rho\sigma}, \qquad \rho = \exp(2\pi i/m).$$

In case $b = 0$, the fibre space F has no singular fibre. Hence F can be represented in the form

$$F = D \times \boldsymbol{C}/G,$$

where G is the group consisting of analytic automorphisms

$$(30) \qquad (\sigma, \zeta) \rightarrow (\sigma, \zeta + n_1\omega(\sigma^m) + n_2), \qquad n_1, n_2 \in \boldsymbol{Z},$$

of $D \times \boldsymbol{C}$. Note that $\omega(\sigma^m)$ is a holomorphic function of $\tau = \sigma^m$, $|\tau| < \epsilon^m$, with $\Im\omega(\sigma^m) > 0$. We denote by $[\sigma, \zeta]$ the point on F corresponding to (σ, ζ).

The generator g of $\mathfrak{G}$ can be written in the form

$$(31) \qquad g\colon [\sigma, \zeta] \rightarrow [\rho\sigma, \zeta + \beta(\sigma)],$$

where $\beta(\sigma)$ is a holomorphic function of σ, $|\sigma| < \epsilon$. Since the order of g is m, we have

$$\sum_{\nu=0}^{m-1} \beta(\rho^\nu\sigma) = h\omega(\sigma^m) + l,$$

where h and l are integers. Hence we infer the existence of a holomorphic function $\gamma(\sigma)$ such that

$$\gamma(\rho\sigma) - \gamma(\sigma) = \beta(\sigma) - (h\omega(\sigma^m) + l)/m.$$

We choose integers a, b, c, d and k such that

$$ad - bc = 1, \qquad kc = h, \qquad kd = l,$$

and introduce a new fibre coordinate

$$\zeta' = (\zeta - \gamma(\sigma))/(c\omega(\sigma^m) + d)$$

on $D \times \boldsymbol{C}$. Correspondingly we replace $\omega(\sigma^m)$ by

$$\omega'(\sigma^m) = (a\omega(\sigma^m) + b)/(c\omega(\sigma^m) + d)$$

and rewrite (30) in the form

$$(\sigma, \zeta') \to (\sigma, \zeta' + n_1\omega'(\sigma^m) + n_2).$$

Then the formula (31) takes the form

$$g\colon [\sigma, \zeta'] \to [\rho\sigma, \zeta' + k/m].$$

Thus, by an appropriate choice of the fibre coordinate ζ, we may assume that g has the form

$$g\colon [\sigma, \zeta] \to [\rho\sigma, \zeta + k/m], \qquad \rho = \exp(2\pi i/m), \tag{32}$$

where the integer k is relatively prime to m. We have

$$S \mid E_a = F/\mathcal{G}.$$

We denote by $((\sigma, \zeta))$ the point on $S \mid E_a$ corresponding to $[\sigma, \zeta]$.

Letting G^* be the group consisting of analytic automorphisms

$$(\tau, \zeta) \to (\tau, \zeta + n_1\omega(\tau) + n_2), \qquad n_1, n_2 \in \boldsymbol{Z}.$$

of $E_a \times \boldsymbol{C}$, we form the quotient space

$$F^* = E_a \times \boldsymbol{C}/G^*$$

and denote by $[\tau, \zeta]^*$ the point on F^* corresponding to (τ, ζ). Note that F^* forms an analytic fibre space of elliptic curves over E_a in an obvious manner. We denote by C^* the fibre of F^* over a.

We write $\Delta' = \Delta - a$, $E_a' = E_a - a$ and $S' = S \mid \Delta'$. In view of (32), the formula

$$\Lambda\colon ((\sigma, \zeta)) \to [\sigma^m, \zeta - (k/2\pi i)\log\sigma]^*, \qquad 0 < |\sigma| < \epsilon, \tag{33}$$

defines a *biholomorphic fibre map* of $S \mid E_a'$ onto $F^* - C^*$. We identify each point $((\sigma, \zeta)) \in S \mid E_a' \subset S'$ with its image $\Lambda((\sigma, \zeta))$ in F^* and form the union

$$S^* = F^* \cup S' \tag{34}$$

of F^* and S'. Obviously *S^* is an analytic fibre space of elliptic curves over Δ of which the fibre C^* over a is regular and $S^* - C^*$ coincides with S'.*

We describe a neighborhood of a singular fibre of type I_n, $n \geqq 1$, of an elliptic surface in the following manner (see Kodaira [9], § 8, (v)). We take a disk $D\colon |\sigma| < \epsilon$ and write $D' = D - 0$. We let $\mathcal{B}$ denote the infinite cyclic group of analytic automorphisms of $D' \times \boldsymbol{C}^*$ generated by the automorphism

$$(\sigma, w) \to (\sigma, w\sigma^n)$$

and form the quotient space

$$W = D \times \boldsymbol{C}^*/\mathfrak{G} = W' \cup \boldsymbol{C}^*, \qquad W' = D' \times \boldsymbol{C}^*/\mathfrak{G}.$$

We denote by $((\sigma, w))$ the point on W corresponding to (σ, w) in $D \times \boldsymbol{C}^*$. We take n copies $W_j = W_j' \cup \boldsymbol{C}_j^*$, $j = 0, 1, 2, \cdots, n-1$ of $W = W' \cup \boldsymbol{C}^*$, identify $((\sigma, w))_k \in W_k'$ with $((\sigma, w\sigma^{j-k}))_j \in W_j'$ and form their union

$$\mathfrak{U}^n = W_0 \cup \cdots \cup W_j \cup \cdots \cup W_{n-1}.$$

In what follows we consider the index j of $W_j, \boldsymbol{C}_j^*, \cdots$ as an element of the cyclic group $\boldsymbol{Z}_n = \boldsymbol{Z}/(n)$ of order n. Thus, for instance, $((\sigma, w))_{-1}$ will denote $((\sigma, w))_{n-1}$. We extend $\mathfrak{U}^n$ to a non-compact complex analytic surface

$$F^n(D) = \mathfrak{U}^n \cup \mathfrak{q}_0 \cup \cdots \cup \mathfrak{q}_j \cup \cdots \cup \mathfrak{q}_{n-1}$$

by adjoining n points $\mathfrak{q}_j$, $j \in \boldsymbol{Z}_n$, in such a manner that

$$\mathfrak{q}_j = \lim_{w \to 0}((0, w))_j = \lim_{w \to \infty}((0, w))_{j-1}.$$

$F^n(D)$ forms an analytic fibre space of elliptic curves over D whose canonical projection is the map: $((\sigma, w))_j \to \sigma$, $\mathfrak{q}_j \to 0$. The fibre $F^n(D)$ over 0 is a singular fibre of type I_n composed of n rational curves

$$\Theta_j = \boldsymbol{C}_j^* \cup \mathfrak{q}_j \cup \mathfrak{q}_{j+1}, \qquad j \in \boldsymbol{Z}_n.$$

Moreover any singular fibre of type I_n, $n \geqq 1$, of an elliptic surface has a neighborhood which is complex analytically equivalent to $F^n(D)$ (see Kodaira [9], Theorem 10.1).

In case $b \geqq 1$, we may assume therefore that $F = F^n(D)$, $n = mb$. The generator g of $\mathfrak{G}$ can be represented in the form

$$g\colon ((\sigma, w))_j \to ((\rho\sigma, w\rho^{j+kb}\beta(\sigma)))_{j+kb}, \qquad \mathfrak{q}_j \to \mathfrak{q}_{j+kb},$$

where $\beta(\sigma)$ is a non-vanishing holomorphic function of σ, $|\sigma| < \epsilon$, and k is an integer which is relatively prime to m. We have

$$\beta(\sigma)\beta(\rho\sigma) \cdots \beta(\rho^{m-1}\sigma) = (-1)^{(m-1)kb}.$$

Hence we infer the existence of a non-vanishing holomorphic function $\gamma(\sigma)$ of σ, $|\sigma| < \epsilon$, such that

$$\beta(\sigma) = \rho^{\nu} c \gamma(\sigma)/\gamma(\rho\sigma), \qquad c = \exp(\pi i(m-1)kb/m),$$

where ν is an integer. We introduce an automorphism

$$s: ((\sigma, w))_j \to ((\sigma, w\gamma(\sigma)))_{j+\nu}, \qquad \mathfrak{q}_j \to \mathfrak{q}_{j+\nu},$$

of F. Then we have

$$sgs^{-1}: ((\sigma, w))_j \to ((\rho\sigma, w\rho^{j+kb}c))_{j+kb}.$$

Therefore we may assume that g has the form

$$g: ((\sigma, w))_j \to ((\rho\sigma, w\rho^{j+kb}c))_{j+kb}, \qquad \mathfrak{q}_j \to \mathfrak{q}_{j+kb}. \tag{35}$$

Writing $((\sigma, w_j))_j$ for $((\sigma, w))_j$ we consider (σ, w_j) as a local coordinate on W_j. Then, since $w_j = w_k\sigma^{j-k}$ for $\sigma \neq 0$, we have

$$d\sigma \wedge dw_j/w_j = d\sigma \wedge dw_k/w_k, \qquad \text{on } W_j \cap W_k.$$

Hence we infer *the existence of a non-vanishing holomorphic 2-form f defined on $F^n(D)$ such that*

$$f = d\sigma \wedge dw_j/w_j, \qquad \textit{on } W_j.$$

We note that

$$\sigma^{m-1}f = \sigma^{m-1}d\sigma \wedge dw_j/w_j$$

represents a holomorphic 2-form defined on $S \mid E_a = F/\mathfrak{G}$.

We form an analytic fibre space

$$F^* = F^b(E_a)$$

of elliptic curves over E_a and denote by C^* the fibre of F^* over a which is of type I_b. We have

$$F^* - C^* = E_a' \times \mathbf{C}^*/\mathfrak{B}^*,$$

where $\mathfrak{B}^*$ is the infinite cyclic group of analytic automorphisms of $E_a' \times \mathbf{C}^*$ generated by the automorphism: $(\tau, w) \to (\tau, w\tau^b)$. We denote by $((\tau, w))^*$ the point on $F^* - C^*$ corresponding to (τ, w). Setting

$$\lambda(\sigma) = \exp(mkb(\log\sigma - \pi i)^2/4\pi i), \qquad \text{for } 0 < |\sigma| < \epsilon,$$

we obtain

$$\lambda(\rho\sigma) = c^{-1}\sigma^{kb}\lambda(\sigma).$$

Therefore, we infer from (35) that the formula

$$\Lambda: ((\sigma, w))_j \to ((\sigma^m, w\sigma^{-j}\lambda(\sigma)))^*, \qquad 0 < |\sigma| < \epsilon, \tag{36}$$

defines a biholomorphic fibre map of $S \mid E_a'$ onto $F^* - \quad {}^*$. We identify $S \mid E_a' \subset S'$ with $F^* - C^* \subset F^*$ by means of Λ and form the union

$$S^* = F^* \cup S'. \tag{37}$$

S^* is an analytic fibre space of elliptic curves over Δ of which the fibre C^* over a is of type I_b and $S^* - C^*$ coincides with S'.

We infer from the above results that

$$S = S \mid E_a \cup (S^* - C^*),$$

where $S \mid E_a' \subset S \mid E_a$ is identified with $F^* - C^* \subset S^* - C^*$ by means of the biholomorphic fibre map Λ. Thus *we obtain S from S^* by replacing the fibre C^* of type I_b by a fibre C_a of type ${}_mI_b$.* We write

$$S = L_a(S^*)$$

and call L_a a *logarithmic transformation.*

Consider a holomorphic 2-form ψ defined on F^*. In case $b = 0$, ψ has the form

$$\psi = \alpha(\tau)\, d\tau \wedge d\zeta,$$

where $\alpha(\tau)$ is a holomorphic function of τ, $|\tau| < \epsilon^m$. Hence we infer from (33) that the logarithmic transformation L_a transforms ψ into the holomorphic 2-form

$$\phi = m\alpha(\sigma^m)\sigma^{m-1} d\sigma \wedge d\zeta \tag{38}$$

on $S \mid E_a$. In case $b \geqq 1$, there exists on F^* a non-vanishing holomorphic 2-form f^* such that

$$f^* = d\tau \wedge dw/w, \qquad \text{on } F^* - C^*,$$

and

$$\psi = \alpha(\tau) f^*,$$

where $\alpha(\tau)$ is a holomorphic function of τ, $|\tau| < \epsilon^m$. Hence we infer from (36) that L_a transforms ψ into the holomorphic 2-form

$$\phi = m\alpha(\sigma^m)\sigma^{m-1} f \tag{39}$$

defined on $S \mid E_a$.

We conclude that *any elliptic surface is obtained from an elliptic surface free from multiple singular fibres by means of a finite number of logarithmic transformations.*

Let S be an analytic fibre space of elliptic curves over a non-singular algebraic curve Δ possessing multiple singular fibres over r points $a_1, \cdots, a_\nu, \cdots, a_r$ on Δ. Letting Ψ denote the canonical projection of S onto Δ, we write the multiple singular fibre over a_ν in the form $m_\nu P_\nu$,

$P_\nu = \Psi^{-1}(a_\nu)$, where m_ν indicates its multiplicity. We denote by $[P_\nu]$ the complex line bundle representing the divisor class of P_ν. Moreover, for any complex line bundle $\mathfrak{f}$ over Δ, we denote by $\Psi^*(\mathfrak{f})$ the complex line bundle over S induced from $\mathfrak{f}$ in an obvious manner.

THEOREM 12. *The canonical bundle K of S has the form*

$$K = \Psi^*(\mathfrak{k} - \mathfrak{f}) + \sum_\nu (m_\nu - 1)[P_\nu], \tag{40}$$

where $\mathfrak{k}$ denotes the canonical bundle of Δ and $\mathfrak{f}$ is a complex line bundle over Δ with

$$c(\mathfrak{f}) = -p_g(S) + q(S) - 1. \tag{41}$$

Proof. Any elliptic surface free from multiple singular fibre is written in the form B^η, $\eta \in H^1(\Delta, \Omega(B^\sharp))$ (see Kodaira [9], Theorem 10.1). Hence, by the above conclusion, S can be represented in the form

$$S = L_{a_r} L_{a_{r-1}} \cdots L_{a_2} L_{a_1}(B^\eta).$$

B is an analytic fibre space of elliptic curves over Δ which possesses a global holomorphic section $o: u \to o(u)$. We identify Δ with the curve $o(\Delta)$ and define $\mathfrak{f}$ to be the normal bundle of $\Delta = o(\Delta)$ in B. Let $\{E_j\}$ be a finite covering of Δ by small circular disks E_j and let $\{f_{jk}\}$ denote a system of transition functions defining the complex line bundle $\mathfrak{k} - \mathfrak{f}$. The canonical bundle of B is induced from $\mathfrak{k} - \mathfrak{f}$ (see [9], Theorem 12.1). This means the existence of a collection of *non-vanishing holomorphic 2-forms* ψ_j defined respectively on $B \mid E_j$ such that

$$\psi_k = f_{jk}\psi_j, \qquad \text{on } B \mid E_j \cap E_k.$$

We obtain B^η from the collection of pieces $B \mid E_j$ of B by identifying $z_j \in B \mid E_j$ with $z_k \in B \mid E_k$ if and only if $z_j = L(\eta_{jk})z_k$ (see [9], Definition 9.2). The fibre preserving analytic automorphism $L(\eta_{jk})$ of $B \mid E_j \cap E_k$ reduces to a translation on each regular fibre of $B \mid E_j \cap E_k$. It follows that $L(\eta_{jk})$ keeps invariant any holomorphic 2-form defined on $B \mid E_j \cap E_k$. Hence, considering each ψ_j as a 2-form defined on $B^\eta \mid E_j$, we obtain

$$\psi_k = f_{jk}\psi_j, \qquad \text{on } B^\eta \mid E_j \cap E_k. \tag{42}$$

This proves that the canonical bundle of B^η is also induced from the complex line bundle $\mathfrak{k} - \mathfrak{f}$.

The transformation $L_{a_r} \cdots L_{a_2} L_{a_1}$ transforms ψ_j into a holomorphic 2-forms ϕ_j on $S \mid E_j$. In case $S \mid E_j$ contains no multiple singular fibre, ϕ_j

does not vanish at each point of $S \mid E_j$. In case $S \mid E_j$ contains a multiple singular fibre, say $m_\nu P_\nu$, ϕ_j has the form (38) or (39). Hence the divisor of ϕ_j is

$$(\phi_j) = (m_\nu - 1)P_\nu.$$

On the other hand, we infer from (42) that

$$\phi_k = f_{jk}\phi_j, \qquad \text{on } S \mid E_j \cap E_k.$$

Consequently we obtain the formula (40).

The Euler number of an elliptic surface is equal to the sum of the Euler numbers of its singular fibres. Therefore $c_2(S)$ is equal to $c_2(B)$, while $c_1^2(S)$ and $c_1^2(B)$ both vanish. Hence, using (3), we get

$$p_g(S) - q(S) = p_g(B) - q(B).$$

On the other hand, by a result of our previous paper ([9], Theorem 12.3), we have

$$c(\mathfrak{f}) = -p_g(B) + q(B) - 1.$$

Consequently we obtain (41).

5. **K3 surfaces.** We shall say that a surface S is a *deformation* of a surface S_0 if there exists a finite sequence of surfaces $S_0, S_1, \cdots, S_k, \cdots, S_n = S$ such that, for each k, S_k and S_{k-1} belong to one and the same complex analytic family $\mathfrak{F}_k$ of surfaces. If S is a deformation of S_0, then S and S_0 belong to one and the same differentiable family of surfaces and S is differentiably homeomorphic to S_0 (see Kodaira and Spencer [14], I, §1). It is clear that *any deformation of* K3 *surface is a* K3 *surface.* In this section we shall prove the following theorem which has been conjectured earlier by A. Weil and independently by A. Andreotti (compare Grauert [5]).

THEOREM 13. *Every* K3 *surface is a deformation of a non-singular quartic surface in a projective* 3*-space.*

i) First we construct a complex analytic family of algebraic elliptic K3 surfaces. Let $\boldsymbol{P}^2$ denote a projective plane on which a system of homogeneous coordinates (x, y, z) is fixed. We take two copies $W_0 = \boldsymbol{P}^2 \times \boldsymbol{C}_0$ and $W_1 = \boldsymbol{P}^2 \times \boldsymbol{C}_1$ of the cartesian product $\boldsymbol{P}^2 \times \boldsymbol{C}$ and form their union

$$W = W_0 \cup W_1$$

by identifying $(x, y, z, u) \in W_0$ with $(x_1, y_1, z_1, u_1) \in W_1$ if and only if

$$(43) \qquad uu_1 = 1, \quad u^4 x_1 = x, \quad u^6 y_1 = y, \quad z_1 = z.$$

Similarly we define

$$\Delta = \boldsymbol{C}_0 \cup \boldsymbol{C}_1,$$

where we identify $u \in \boldsymbol{C}_0$ with $u_1 \in \boldsymbol{C}_1$ if and only if $uu_1 = 1$. Obviously W is an algebraic manifold and

$$\Psi : (x, y, z, u) \to u, \qquad (x_1, y_1, z_1, u_1) \to u_1$$

is a holomorphic map of W onto the projective line Δ. Thus W is a complex analytic fibre bundle of projective planes over Δ. For the sake of simplicity we denote by $u = \infty$ the point $u_1 = 0$ on Δ.

For a point $\tau = (\tau_0, \tau_1, \cdots, \tau_8, \sigma_1, \cdots, \sigma_{12})$ in the space $\boldsymbol{C}^{21}$ of 21 complex variables we set

$$g(u) = \tau_0 \prod_{\nu=1}^{8} (u - \tau_\nu), \qquad h(u) = \prod_{\nu=1}^{12} (u - \sigma_\nu),$$

and we define B_τ to be the algebraic surface in W determined by the pair of equations

$$\begin{cases} y^2 z - 4x^3 + g(u) x z^2 + h(u) z^3 = 0, \\ y_1^2 z_1 - 4x_1^3 + u_1^8 g(1/u_1) x_1 z_1^2 + u_1^{12} h(1/u_1) z_1^3 = 0. \end{cases}$$

Note that these two equations are equivalent in the intersection $W_0 \cap W_1$. We set

$$\mathfrak{J}_\tau = \mathfrak{J}_\tau(u) = g(u)^3 / (g(u)^3 - 27 h(u)^2). \tag{44}$$

Assume that τ satisfies the following three conditions:

1) $\tau_0 \neq 0$, $\tau_0^3 \neq 27$;

2) if $g(\sigma_\lambda) = 0$, then $\sigma_\nu \neq \sigma_\lambda$ for $\nu \neq \lambda$;

3) the meromorphic function $\mathfrak{J}_\tau$ has no multiple pole.

Then we infer by an elementary consideration that B_τ is non-singular and that the intersection

$$C(u) = B_\tau \cdot \Psi^{-1}(u)$$

is a plane cubic curve with multiplicity 1. Thus B_τ *is an analytic fibre space of elliptic curves over* Δ *of which the canonical projection is the restriction* Ψ_τ *of* Ψ *to* B_τ. Since the equation of the fibre $C(u)$, $u \neq \infty$, is

$$y^2 z - 4x^3 + g(u) x z^2 + h(u) z^3 = 0,$$

the functional invariant of B_τ *is* $\mathfrak{J}_\tau$ (see Kodaira [9], § 7).

In order to examine the singular fibres of B_T we may assume that

$$\begin{aligned} &\sigma_\nu = \tau_\nu, && \text{for } \nu = 1, 2, \cdots, r, \\ &\sigma_\nu \neq \tau_\lambda, && \text{for } \nu = r+1, \cdots, 12. \end{aligned}$$

Let $a_1, \cdots, a_\rho, \cdots, a_j$ denote the poles of $\mathfrak{J}_T$ which are, by hypothesis, simple. Then we have

$$g(u)^3 - 27h(u)^2 = (\tau_0^3 - 27) \prod_{\nu=1}^{r} (u - \tau_\nu)^2 \prod_{\rho=1}^{j} (u - a_\rho), \tag{45}$$

where $\tau_\nu \neq \tau_\lambda$ for $1 \leqq \nu < \lambda \leqq r$. Hence we infer that, if $u \neq \tau_1, \cdots, \tau_r, a_1, \cdots, a_j, \infty$, then $C(u)$ is a non-singular elliptic curve and that $C(a_\rho)$ is a rational curve with one ordinary double point. Each fibre $C(\tau_\nu)$, $1 \leqq \nu \leqq r$, is a rational curve with one cuspidal point defined by the equation: $y^2z - 4x^3 = 0$. The fibre $C(\infty)$ is a non-singular elliptic curve, since it is defined by the equation

$$y_1^2 z_1 - 4x_1^3 + \tau_0 x_1 z_1^2 + z_1^3 = 0, \qquad \tau_0^3 \neq 27.$$

Thus *the elliptic surface B_T has j singular fibres $C(a_1), \cdots, C(a_j)$ of type I_1 and r singular fibres $C(\tau_1), \cdots, C(\tau_r)$ of type II.* It follows from (45) that

$$j + 2r = 24. \tag{46}$$

The equation $x = z = x_1 = z_1 = 0$ determines a projective line P in W and Ψ maps P biholomorphically onto Δ. Thus P represents a holomorphic section of the complex analytic fibre bundle W over Δ. We identify Δ with P by means of the biholomorphic map Ψ. It is obvious that $\Delta = P \subset B_T$. Hence Δ represents a *global holomorphic section* of the elliptic surface B_T.

We set

$$\begin{gathered} f = y^2 z - 4x^3 + g(u) x z^2 + h(u) z^3, \\ f_1 = y_1^2 z_1 - 4x_1^3 + u_1^8 g(1/u_1) x_1 z_1^2 + u_1^{12} h(1/u_1) z_1^3 \end{gathered}$$

and we write $f_x, f_{1x_1}, f_y, \cdots$ for $\partial f/\partial x, \partial f_1/\partial x_1, \partial f/\partial y, \cdots$, respectively. Then the formula

$$\begin{aligned} \psi &= du \wedge (z dx - x dz)/f_y = du_1 \wedge (x_1 dz_1 - z_1 dx_1)/f_{1y_1} \\ &= du \wedge (x dy - y dx)/f_z = du_1 \wedge (y_1 dx_1 - x_1 dy_1)/f_{1z_1} \\ &= du \wedge (y dz - z dy)/f_x = du_1 \wedge (z_1 dy_1 - y_1 dz_1)/f_{1x_1} \end{aligned}$$

defines a *non-vanishing holomorphic 2-form* ψ on B_T. Hence the canonical bundle of B_T is trivial and the geometric genus p_g of B_T is 1. Since the Euler number c_2 of B_T is equal to the sum of the Euler numbers of the singular fibres of B_τ, it follows from (46) that c_2 is equal to 24. Combining these results with (3) we infer that the irregularity q of B_τ vanishes. Thus we conclude that *B_τ is a* K3 *surface.*

We denote by N the *connected open subset* of $\boldsymbol{C}^{21}$ which consists of all points τ satisfying the conditions 1), 2, 3) and define $\mathfrak{B}$ to be the submanifold of $W\times N$ determined by the equations

$$\begin{cases} y^2z-4x^3+g(u,\tau)xz^2+h(u,\tau)z^3=0, \\ y_1^2z_1-4x_1^3+u_1^8g(1/u_1,\tau)x_1z_1^2+u_1^{12}h(1/u_1,\tau)z_1^3=0, \end{cases}$$

where we write $g(u,\tau)$ and $h(u,\tau)$ for $g(u)$ and $h(u)$, respectively, in order to make explicit that they are polynomials in u and τ. Clearly $\mathfrak{B}$ cuts out on each submanifold $W\times\tau$ the surface $B_\tau\times\tau$. Thus we obtain a *complex analytic family $\mathfrak{B}$ consisting of algebraic elliptic* K3 *surfaces B_τ, $\tau\in N$.*

ii) In this subsection we assume that *S is a* K3 *surface.* The geometric genus p_g of S is equal to 1. Hence, using the formula (3) and Theorem 3, we get

(47) $$b_1=h^{1,0}=0,\quad b^+=3,\quad b^-=19,\quad b_2=22,\quad c_2=24.$$

Combining these results with the identities

$$h^{r,\nu}=h^{2-r,2-\nu},\qquad \sum(-1)^{r+\nu}h^{r,\nu}=c_2,$$

we obtain

(48) $$h^{1,1}=20,\qquad h^{1,2}=0.$$

Let $\{\Gamma_1,\cdots,\Gamma_j,\cdots,\Gamma_{22}\}$ be a Betti base of 2-cycles on S. For any cohomology class $c\in H^2(S,\boldsymbol{Z})$, we denote by $c(\Gamma_j)$ the value of c on Γ_j. We choose cohomology classes $e_j\in H^2(S,\boldsymbol{Z})$ such that

$$e_j(\Gamma_k)=\delta_{jk},\qquad\qquad \text{for } j,k=1,2,\cdots,22,$$

and define

$$A(x,y)=\sum_{j,k=1}^{22}a_{jk}x_jy_k,\qquad\qquad a_{jk}=e_je_k,$$

where the x_j and the y_k are indeterminates. Note that any element c of $H^2(S,\boldsymbol{Z})$ can be written in the form

$$c=\sum m_je_j+\text{torsion element},\qquad\qquad m_j=c(\Gamma_j),$$

and

(49) $$c^2=A(m,m).$$

The natural injection: $\boldsymbol{Z}\to\boldsymbol{R}$ induces a homomorphism

$$\iota\colon H^2(S,\boldsymbol{Z})\to H^2(S,\boldsymbol{R}).$$

We represent each cohomology class ιe_j by a real d-closed 2-form η_j. We have

$$(50) \qquad (\eta_j\eta_k) = a_{jk}.$$

There exists on S a non-vanishing holomorphic 2-form ψ. Following A. Weil and A. Andreotti, *we associate with* S *the point*

$$(51) \qquad \lambda = \lambda(S) = (\lambda_1, \cdots, \lambda_j, \cdots, \lambda_{22}), \qquad \lambda_j = \int_{\Gamma_j} \psi,$$

in a projective space P^{21} *of dimension* 21. We have

$$\psi = \sum \lambda_j \eta_j + dv,$$

where v is a 1-form on S. Hence, using (50), we obtain

$$A(\lambda, \lambda) = \int_S \psi \wedge \psi = 0.$$

Thus λ *is on the hypersurface* M *in* P^{21} *defined by the quadric equation* $A(x,x) = 0$. This result is due to A. Weil and A. Andreotti.

THEOREM 14. *A cohomology class* $c \in H^2(S, Z)$ *is the Chern class of a complex line bundle* F *over* S *if and only if the point*

$$m = (m_1, \cdots, m_j, \cdots, m_{22}), \qquad m_j = c(\Gamma_j),$$

satisfies the linear equation

$$(52) \qquad A(\lambda, m) = 0, \qquad \lambda = \lambda(S).$$

Proof. In view of the exact sequence

$$0 \to H^1(S, \mathcal{O}^*) \to H^2(S, Z) \xrightarrow{\kappa} H^2(S, \mathcal{O}) \to \cdots,$$

there exists a complex line bundle F over S with $c(F) = c$ if and only if κc vanishes. Let

$$\eta = \sum m_j \eta_j$$

and let $\eta^{(2)}$ denote the component of η of type $(0,2)$. The d-closed 2-form η represents the cohomology class ιc. Hence the $\bar{\partial}$-closed $(0,2)$-form $\eta^{(2)}$ represents the cohomology class κc by means of the Dolbeault isomorphism. The $\bar{\partial}$-closed $(0,2)$-form $\bar{\psi}$ is not $\bar{\partial}$-cohomologous to zero. In fact, the existence of a $(0,1)$-form σ with $\bar{\partial}\sigma = \bar{\psi}$ would imply that

$$\int_S \psi \wedge \bar{\psi} = \int_S d(\psi \wedge \sigma) = 0.$$

Moreover the dimension of $H^2(S, \mathfrak{G})$ is equal to $p_g = 1$. Consequently we obtain

$$\eta^{(2)} = \beta\bar{\psi} + \bar{\partial}\sigma,$$

where β is a constant and σ denotes a $(0,1)$-form. Evidently κc vanishes if and only if $\beta = 0$. Now, using (50), we obtain

$$\beta \int_S \psi \wedge \bar{\psi} = \int_S \psi \wedge \eta^{(2)} = \int_S \psi \wedge \eta = A(\lambda, m).$$

Hence we conclude that there exists a complex line bundle F over S with $c(F) = c$ if and only if $A(\lambda, m)$ vanishes, q. e. d.

A point in a projective space is said to be *rational* if its homogeneous coordinates are rational numbers.

THEOREM 15. *If the linear equation*

$$A(\lambda, m) = 0, \qquad \lambda = \lambda(S), \tag{53}$$

has one and only one rational solution m in $\mathbf{P}^{21}$ and if that solution m satisfies the quadric equation

$$A(m, m) = 0, \tag{54}$$

then S is an analytic fibre space of elliptic curves over a projective line Δ whose singular fibres are either of type I_1 or of type II.

Proof. We choose the homogeneous coordinates m_j of m such that $m_1, \cdots, m_j, \cdots, m_{22}$ are rational integers having no common divisors and let

$$e = \sum m_j e_j.$$

In view of Theorem 14, the uniqueness of the rational solution m of (53) implies that the Chern class of any complex line bundle F over S has the form

$$c(F) = ne + \text{torsion element},$$

where n denotes an integer. Hence we get

$$c(F)^2 = n^2 A(m, m) = 0. \tag{55}$$

This proves that S is non-algebraic.

By Theorem 14 there exists a complex line bundle E over S with $c(E) = e$. Since the canonical bundle of S is trivial, we obtain from (4) the inequality

$$\dim |E| + \dim |-E| \geqq 0.$$

This implies that either $|E|$ or $|-E|$ is non-empty. We suppose that $|E|$ is non-empty. (In case $|E|$ is empty we replace E by $-E$.) Then, since $c(E)=e\neq 0$, $|E|$ contains a positive divisor D and therefore

$$\dim |E| \geqq -\dim |-D| = 1.$$

This proves the existence of non-constant meromorphic functions on S, while S is non-algebraic. Consequently S is an elliptic surface.

Since $h^{1,0}=0$, the base curve Δ of S is a projective line. In view of the formula (40), the triviality of the canonical bundle of S implies that S has no multiple singular fibres. Moreover the formula (55) shows that S contains no curve C with $(C^2)\neq 0$. Hence each singular fibre of S is either of type I_1 or of type II (see Kodaira [9], Theorem 6.2).

iii) We shall examine the structure of an elliptic K3 surface S whose singular fibres are either of type I_1 or of type II. The base curve Δ is a projective line. We fix a non-homogeneous coordinate on Δ and indicate a point on Δ by its non-homogeneous coordinate u. We denote by $\mathfrak{J}=\mathfrak{J}(u)$ the *functional invariant* of S and by G the *homological invariant* of S (see [9], § 7). Then we have

$$S = B^{\eta}, \qquad \eta\in H^1(\Delta, \Omega(B_0{}^{\#})),$$

where B is the basic member of the family $\mathfrak{F}(\mathfrak{J}, G)$ (see [9], Theorem 11.1 and § 8). For the sake of simplicity we assume that $\mathfrak{J}(\infty)\neq 0,1,\infty$. Let $C(u)$ denote the fibre of S over u. There are certain relations between the structure of $C(u)$ and the behavior of $\mathfrak{J}$ at u (see [9], § 8 and Table I in § 9). We infer from these relations that the meromorphic function $\mathfrak{J}$ has no multiple pole, that $C(u)$ is a singular fibre of type I_1 if and only if u is a simple pole of $\mathfrak{J}$ and that $C(u)$ is a singular fibre of type II if and only if u is a root of $\mathfrak{J}$ of multiplicity $\equiv 1 \bmod (3)$. If u is a root of $\mathfrak{J}$ and if $C(u)$ is a regular fibre, then its multiplicity is divisible by 3. The multiplicity of each root of $\mathfrak{J}-1$ is even. Moreover, letting j and r denote respectively the numbers of the singular fibres of S of type I_1 and of type II, we have

$$j+2r=c_2=24.$$

Thus we infer that

$$\mathfrak{J}/(\mathfrak{J}-1)=\alpha\prod_{\nu=1}^{r}(u-\tau_\nu)\prod_{\nu=r+1}^{8}(u-\tau_\nu)^3/\prod_{\mu=r+1}^{12}(u-\sigma_\mu)^2, \tag{56}$$

where $\alpha\neq 0$, $\tau_\nu\neq\sigma_\mu$, and $\tau_\nu\neq\tau_\rho$ for $1\leqq \nu<\rho\leqq r$. We denote the (simple) poles of $\mathfrak{J}$ by $a_1, a_2, \cdots, a_j$ and let $a_{j+\nu}=\tau_\nu$ for $\nu=1,2,\cdots,r$. We conclude that S has singular fibres $C(a_1), \cdots, C(a_j)$ of type I_1 and $C(a_{j+1}), \cdots, C(a_{j+r})$ of type II.

LEMMA 2. *The homological invariant G of S is uniquely determined by $\mathcal{J}$.*

Proof. We denote by $(\omega(u), 1)$, $\Im\omega(u) > 0$, the periods of the elliptic curve $C(u)$, $u \neq a_\rho$. $\omega(u)$ is a multi-valued holomorphic function on the punctured sphere $\Delta' = \Delta - \{a_\rho\}$ and $\mathcal{J}(u) = J(\omega(u))$, where J denotes the elliptic modular function (see [9], § 7). We fix a point b on Δ' and represent each element of the fundamental group $\pi_1(\Delta')$ by a closed continuous curve β on Δ' starting and ending at b. The analytic continuation along β induces a modular transformation

$$\omega(u) \rightarrow R(\beta)\omega(u) = (a_\beta\omega(u) + b_\beta)/(c_\beta\omega(u) + d_\beta),$$

where a_β, b_β, c_β, d_β are rational integers and $a_\beta d_\beta - b_\beta c_\beta = 1$. The homological invariant G is determined by a representation of $\pi_1(\Delta')$ of the form

$$(R) : \beta \rightarrow (R)(\beta) = \pm \begin{pmatrix} a_\beta & b_\beta \\ c_\beta & d_\beta \end{pmatrix} \tag{57}$$

(see [9], § 7). To prove the uniqueness of G it suffices to show the uniqueness of the representation (R).

For each point a_ρ we choose a closed continuous curve α_ρ on Δ' starting and ending at b which is homotopic on Δ' to a positively oriented small circle around a_ρ. Obviously the closed curves $\alpha_1, \cdots, \alpha_\rho, \cdots, \alpha_{j+r}$ generate the fundamental group $\pi_1(\Delta')$. For each generator α_ρ the matrix $(R)(\alpha_\rho)$ is equivalent to $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$ according as the corresponding singular fibre $C(a_\rho)$ is of type I_1 or of type II (see [9], § 9, Table I), while $(R)(\alpha_\rho)$ is determined by (57) uniquely up to a factor ± 1. Hence $(R)(\alpha_\rho)$ is uniquely determined. It follows that the representation (R) is unique, q. e. d.

Let

$$\tau = (\tau_0, \tau_1, \cdots, \tau_8, \tau_1, \cdots, \tau_r, \sigma_{r+1}, \cdots, \sigma_{12}), \qquad \tau_0{}^8 = 27\alpha.$$

Then, comparing (56) with (44), we infer that the functional invariant $\mathcal{J}_\tau$ of the elliptic surface B_τ coincides with $\mathcal{J}$. Hence, by the above Lemma 2, the homological invariant of B_τ coincides with G. Moreover B_τ possesses a global holomorphic section over Δ. Consequently B_τ coincides with the *basic member* B of the family $\mathcal{F}(\mathcal{J}, G)$ (see [9], Theorem 10.2).

LEMMA 3. *The cohomology group $H^2(\Delta, G)$ vanishes.*

Proof. In order to make explicit that G is the homological invariant of B_τ, we write G_τ for G. We have

$$H^2(\Delta, G_\tau) \cong \boldsymbol{Z}^2/M_\tau,$$

where M_τ is the submodule of $\boldsymbol{Z}^2 = \boldsymbol{Z} \oplus \boldsymbol{Z}$ generated by the elements

$$(m,n)(R)(\beta) - (m,n), \qquad (m,n) \in \boldsymbol{Z}^2, \beta \in \pi_1(\Delta').$$

We infer readily that M_τ depends only on the number r of singular fibres of type II of B_τ. Moreover M_τ is lower semi-continuous in τ in the sense that, for each point $\tau \in N$, there exists a positive number ϵ such that $M_\tau \subseteq M_\zeta$ for $|\zeta - \tau| < \epsilon$. To prove the vanishing of $\boldsymbol{Z}^2/M_\tau$, it suffices therefore to consider the case in which $r \geqq 1$. Now, if $r \geqq 1$, then, assuming the normal form $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$ for $(R)(\alpha_{j+1})$, we obtain

$$(m,n)(R)(\alpha_{j+1}) - (m,n) = (-n, m-n).$$

This proves that $M_\tau = \boldsymbol{Z}^2$, q. e. d.

We employ the notation of our previous paper [9]. By the above Lemma 3, the exact sequence (11.7) in [9] is reduced to

$$\cdots \to H^1(\Delta, \Omega(\mathfrak{f})) \xrightarrow{h^*} H^1(\Delta, \Omega(B_0{}^{\#})) \to 0.$$

Hence we infer that $S = B^{h^*(s)}$, $s \in H^1(\Delta, \Omega(\mathfrak{f}))$ (see [9], Theorem 11.1). Thus we obtain the following

THEOREM 16. *Any elliptic* K3 *surface whose singular fibres are of type I_1 or of type II can be represented in the form*

$$S = B^{h^*(s)}, \qquad s \in H^1(\Delta, \Omega(\mathfrak{f})), B \in \mathfrak{B}.$$

We note that the collection of the elliptic surfaces $B^{h^*(s)}$, $s \in H^1(\Delta, \Omega(\mathfrak{f}))$, forms a complex analytic family (see [9], Theorem 11.3). *The elliptic* K3 *surface S is therefore a deformation of $B \in \mathfrak{B}$.*

iv) We consider an *arbitrary* K3 surface S and write the non-vanishing holomorphic 2-form ψ on S explicitly in the form

$$\psi = \tfrac{1}{2} \sum_{\alpha,\beta=1}^{2} \psi_{\alpha\beta} dz_\alpha \wedge dz_\beta, \qquad \psi_{\beta\alpha} = -\psi_{\alpha\beta}.$$

The mapping

(58) $$\sum_\alpha \theta_\alpha \partial/\partial z_\alpha \to \sum_{\alpha,\beta} \theta_\alpha \psi_{\alpha\beta} dz_\beta$$

induces an isomorphism

(59) $$\Theta \cong \Omega^1.$$

Combining this with (48) we obtain

$$\dim H^1(S, \Theta) = 20, \qquad \dim H^2(S, \Theta) = 0. \tag{60}$$

We denote by $t = (t_1, t_2, \cdots, t_{20})$ a point in C^{20} and by U_δ a spherical domain consisting of all points t, $|t| < \delta$, where δ is a positive number. In view of (60) we infer from a theorem of Kodaira, Nirenberg and Spencer [13] the existence of a complex analytic family $\mathcal{F}$ of surfaces S_t, $t \in U_\delta$, such that $S_0 = S$ and such that the map

$$\rho_0 : \partial/\partial t \rightarrow (\partial S_t/\partial t)_{t=0}$$

maps the tangent space $T_0(U_\delta)$ isomorphically onto $H^1(S, \Theta)$, provided that δ is sufficiently small, where the symbol $\partial S_t/\partial t$ denotes the infinitesimal deformaiton of S_t along $\partial/\partial t$. Let Θ_t denote the sheaf over S_t of germs of holomorphic vector fields. Since each surface S_t is a K3 surface, the dimension of $H^1(S_t, \Theta_t)$ is independent of t. Hence we infer from a stability theorem (see Kodaira and Spencer [14], III, §1) that $\rho_t : \partial/\partial t \rightarrow \partial S_t/\partial t$ maps $T_t(U_\delta)$ isomorphically onto $H^1(S_t, \Theta_t)$. Combining this with a theorem of completeness (Kodaira and Spencer [15]) we conclude that the family $\mathcal{F}$ *is effectively parametrized and complete.*

We may assume that the surface S_t, $t \in U_\delta$, have one and the same underlying differentiable manifold X. We identify X with S and employ local complex coordinates (z_1, z_2) as local differentiable coordinates on X. Moreover we choose local complex coordinates (w_1, w_2) on the surface S_t such that $w_1 = w_1(z, t)$ and $w_2 = w_2(z, t)$ are differentiable functions of z_1 and z_2 which depend holomorphically on t and satisfy the boundary conditions $w_1(z, 0) = z_1$, $w_2(z, 0) = z_2$ (see Kodaira, Nirenberg and Spencer [13]). There exists on each surface S_t a non-vanishing holomorphic 2-form $\psi(t)$. We may assume that $\psi(t)$ *depends holomorphically on* t. This means that, when $\psi(t)$ is written in the form

$$\psi(t) = \tfrac{1}{2} \sum_{\alpha,\beta=1}^{2} \psi_{\alpha\beta}(w, t)\, dw_\alpha(z, t) \wedge dw_\beta(z, t), \tag{61}$$

the coefficients $\psi_{\alpha\beta}(w, t)$ are holomorphic functions in w_1, w_2 and t. Supposing that the 2-cycles Γ_j, $j = 1, 2, \cdots, 22$, are fixed on X, we set

$$\lambda_j(t) = \int_{\Gamma_j} \psi(t).$$

The periods $\lambda_j(t)$ are holomorphic functions of t and satisfy the quadric equation

$$A(\lambda(t), \lambda(t)) = \sum a_{jk} \lambda_j(t) \lambda_k(t) = 0.$$

Thus

$$t \to \lambda(t) = \lambda(S_t) = (\lambda_1(t), \cdots, \lambda_j(t), \cdots, \lambda_{22}(t))$$

is a holomorphic map of U_δ into M. The following theorem is due to A. Weil and A. Andreotti.

THEOREM 17. *The map: $t \to \lambda(t)$ maps U_δ biholomorphically onto an open subset of M, provided that δ is sufficiently small.*

Proof. We take a tangent vector $\partial/\partial t \in T_0(U_\delta)$ and, for any holomorphic function $f(t)$ of t, we write

$$f'(0) = (\partial f(t)/\partial t)_{t=0}.$$

For our purpose it suffices to show that the simultaneous equations

$$(62) \qquad \lambda_j(0)\lambda'_k(0) - \lambda_k(0)\lambda'_j(0) = 0, \qquad j, k = 1, 2, \cdots, 22,$$

hold only if $\partial/\partial t$ vanishes. Assume that (62) hold. Then we have

$$\lambda'_j(0) = a\lambda_j(0), \qquad j = 1, 2, \cdots, 22,$$

where a is a constant. Hence we get

$$(63) \qquad \psi'(0) - a\psi(0) = dv,$$

where v is a 1-form on S. From (61) we obtain

$$(64) \qquad \psi'(0) = \sum \psi_{\alpha\beta}(z, 0)\bar{\partial} w'_\alpha(z, 0) \wedge dz_\beta + (2, 0)\text{-form}.$$

We write v as a sum

$$v = v^{(1)} + v^{(0)}$$

of a (0, 1)-form $v^{(1)}$ and a (1, 0)-form $v^{(0)}$. We then infer from (63) and (64) that

$$\bar{\partial} v^{(1)} = 0.$$

Since $H^1(S, \mathcal{O})$ vanishes, this implies that

$$v^{(1)} = \bar{\partial}\mu,$$

where μ is a differentiable function on S. Substituting $\bar{\partial}\mu$ for $v^{(1)}$ in (63) we get

$$\psi'(0) - a\psi(0) = \bar{\partial}(v^{(0)} - \partial\mu) + \partial v^{(0)},$$

Comparing this with (64) we obtain

$$(65) \qquad \sum \psi_{\alpha\beta}(z, 0)\bar{\partial} w'_\alpha(z, 0) \wedge dz_\beta = \bar{\partial}(v^{(0)} - \partial\mu).$$

The mapping (58) induces an isomorphism of the $\bar{\partial}$-cohomology group of vector $(0,1)$-forms onto the $\bar{\partial}$-cohomology group of $(1,1)$-forms. The infinitesimal deformation $\rho_0(\partial/\partial t)$ is represented by the $\bar{\partial}$-cohomology class of the $\bar{\partial}$-closed vector $(0,1)$-form

$$\sum_{\alpha=1}^{2} \bar{\partial} w'_\alpha(z,0)\partial/\partial z_\alpha$$

(see Kodaira, Nirenberg and Spencer [13]). The mapping (58) maps this vector $(0,1)$-form onto the $\bar{\partial}$-closed $(1,1)$-form

$$\sum \psi_{\alpha\beta}(z,0)\bar{\partial} w'_\alpha(z,0) \wedge dz_\beta.$$

Hence it follows from (65) that $\rho_0(\partial/\partial t)$ vanishes. This proves the vanishing of $\partial/\partial t$, q. e. d.

Now we shall prove that, for any positive number $\epsilon \leqq \delta$, at least one of the deformations S_t, $|t| < \epsilon$, of S is an elliptic K3 surface whose singular fibres are of type I_1 or of type II. For this purpose it suffices to show the existence of a point $t \in U_\epsilon$ such that $\lambda(t)$ satisfies the hypotheses of Theorem 15. Let $\boldsymbol{Q}$ denote the field of rational numbers.

LEMMA 4. *If a point $\mu \in M$ is rational over $\boldsymbol{Q}(i)$, $i = \sqrt{-1}$, then there exists a rational point $m \in \boldsymbol{P}^{21}$ satisfying the equations*

$$A(\mu, m) = A(m, m) = 0. \tag{66}$$

Proof. It suffices to consider the case where μ is not rational. Considering μ as a vector in $\boldsymbol{C}^{22}$, we write $\mu = r + is$, where r and s are rational vectors which are not zero. Since $A(\mu,\mu)$ vanishes, we have

$$A(r,s) = 0, \qquad A(r,r) = A(s,s).$$

If $A(r,r) = A(s,s) = 0$, then, setting $m = r$, we obtain (66).

If $A(r,r) = A(s,s) \neq 0$, then we can choose rational vectors r_j, $j = 1, 2, \cdots, 20$, such that

$$A(r,r_j) = A(s,r_j) = 0, \quad A(r_j,r_j) \neq 0, \quad A(r_k,r_j) = 0 \text{ for } k \neq j.$$

Since $b^+ = 3$ and $b^- = 19$, the signature of the quadric form

$$\sum_{j=1}^{20} A(r_j,r_j)x_j^2$$

is either $(3,16)$ or $(1,19)$. Hence the quadric equation

$$\sum_{j=1}^{20} A(r_j,r_j)m_j^2 = 0$$

admits non-trivial rational solutions $m_1, \cdots, m_j, \cdots, m_{20}$ (see Skolem [17], Kap. III, § 1). Consequently, setting $m = \sum m_j r_j$, we obtain a rational vector $m \neq 0$ which satisfies the equations (66), q. e. d.

For each rational point n in P^{21} we define Ξ_n to be the subvariety of M determined by the equations

$$A(x, n) = A(x, x) = 0.$$

M contains an everywhere dense subset of points which are rational over $Q(i)$ (see Skolem [17], Kap. III, § 1). Hence the open subset $\lambda(U_\epsilon)$ of M contains a point μ which is rational over $Q(i)$. With the aid of the above Lemma 4 we find a rational point m in M such that $\mu \in \Xi_m$. Then $\lambda(U_\epsilon) \cap \Xi_m$ is a non-empty open subset of Ξ_m, while $\Xi_m \cap \Xi_n$ is a proper subvariety of Ξ_m for any rational point $n \neq m$ in P^{21}. Consequently we can find a point $\lambda(t) \in \lambda(U_\epsilon) \cap \Xi_m$ such that $\lambda(t) \notin \Xi_n$ for all rational points $n \neq m$. It is clear that this point $\lambda(t)$ satisfies the hypotheses of Theorem 15.

Thus we see that there exists a point $t \in U_\epsilon$ such that S_t is an elliptic K3 surface whose singular fibres are of type I_1 or of type II. Applying Theorem 16 to this surface S_t, we obtain the following

THEOREM 18. *For any* K3 *surface* S, *there exists an arbitrarily small deformation* S_t *of* S *which is an elliptic* K3 *surface of the form*

$$S_t = B^{\eta^*(s)}, \qquad s \in H^1(\Delta, \Omega(\mathfrak{f})), B \in \mathscr{B}.$$

This theorem shows that every K3 surface is a deformation of a member of the complex analytic family $\mathscr{B}$. Any non-singular quartic surface in a projective 3-space is a deformation of a member of $\mathscr{B}$, since it is a K3 surface. Consequently every K3 surface is a deformation of a non-singular quartic surface in a projective 3-space. This completes the proof of Theorem 13.

6. Surfaces with trivial canonical bundles.

Let S be a surface with a trivial canonical bundle. Since the first Chern class c_1 of S vanishes, it follows from a result of Atiyah and Hirzebruch [1] that the Todd genus of S is even, while the gemetric genus p_g of S is equal to 1. Hence, by (3), *the irregularity* q *of* S *is even.* Moreover the triviality of the canonical bundle implies that S contains no exceptional curve.

A) In case b_1 is even, it follows from (13) that

$$8q + b^- = 19.$$

Hence q is equal to either 0 or 2.

If $q = 0$, then S is a K3 surface.

If $q=2$, then there exist on S two linearly independent d-closed holomorphic 1-forms ϕ_1 and ϕ_2 and there is defined a holomorphic map

$$\Phi: z \to \Phi(z) = \left(\int^z \phi_1, \int^z \phi_2\right) \pmod{\mathfrak{D}}$$

of S into the Albanese variety $\mathfrak{A}$ attached to S (see (15)).

To verify that $\phi_1 \wedge \phi_2$ does not vanish, we suppose that $\phi_1 \wedge \phi_2$ vanishes identically. Then the image $\Phi(S)$ is an irreducible curve in $\mathfrak{A}$. Hence we infer that there exists a holomorphic map Ψ of S onto a non-singular algebraic curve Δ of genus 2 such that the inverse image $C_u = \Psi^{-1}(u)$ of any general point $u \in \Delta$ is an irreducible non-singular curve. Since the canonical bundle K of S is trivial, we have

$$2\pi(C_u) - 2 = (C_u^2) + (KC_u) = 0.$$

Thus S is an elliptic surface and therefore, by (40),

$$\Psi^*(\mathfrak{k} - \mathfrak{f}) + \sum (m_\nu - 1)[P_\nu] = K = 0.$$

It follows that

$$c(\mathfrak{f}) = c(\mathfrak{k}) = 2.$$

This contradicts the formula (41).

In view of the triviality of the canonical bundle of S, the holomorphic 2-form $\phi_1 \wedge \phi_2$ does not vanish at each point of S. Hence Φ is a locally biholomorphic map of S onto $\mathfrak{A}$ and therefore S is an unramified covering surface of $\mathfrak{A}$. Consequently S is a complex torus.

B) In case b_1 is odd, it follows from (13) that

$$8q + b^- = 18,$$

while, by Theorem 3, $2q = b_1 + 1$. Hence q is 2 and b_1 is 3. In view of Theorems 4 and 11, S is therefore an elliptic surface. Let Δ be the base curve of S and let Ψ denote the canonical projection of S onto Δ. By (40) we have

$$\Psi^*(\mathfrak{k} - \mathfrak{f}) + \sum (m_\nu - 1)[P_\nu] = K = 0.$$

Hence we infer that S has no multiple singular fibre and that $\mathfrak{k} = \mathfrak{f}$. Since S can not be deformed into an algebraic surface, S is a complex analytic fibre bundle of elliptic curves over Δ whose structure group is the group of translations (see [9], § 11). Using (41) we get

$$c(\mathfrak{k}) = c(\mathfrak{f}) = -p_g + q - 1 = 0.$$

Hence the base curve Δ is elliptic.

We fix a point o on S, denote by C the fibre of S passing through o and choose circles γ_1 and γ_2 on C which pass through o and form a base of 1-cycles on C. Moreover we choose closed continuous curves γ_3 and γ_4 on S starting and ending at o such that $\Psi(\gamma_3)$ and $\Psi(\gamma_4)$ form a base of 1-cycles on Δ and define g_1, g_2, g_3, g_4 to be the elements of the fundamental group $\pi_1(S) = \pi_1(S, o)$ determined respectively by $\gamma_1, \gamma_2, \gamma_3, \gamma_4$. Clearly the projection Ψ induces a homomorphism of $\pi_1(S)$ onto the fundamental group $\pi_1(\Delta)$ of Δ whose kernel κ is the subgroup of $\pi_1(S)$ generated by g_1 and g_2. It follows that $\pi_1(S)$ is generated by g_1, g_2, g_3, g_4. The commutator $g_3 g_4 g_3^{-1} g_4^{-1}$ belongs to κ. Hence, by an appropriate choice of γ_1 and γ_2, we obtain

$$g_3 g_4 = g_2^m g_4 g_3, \tag{67}$$

where m is an integer.

Let ϕ denote the holomorphic 1-form on S induced from a non-vanishing holomorphic 1-form on Δ. We infer in the same manner as in Section 3, B ii) the existence of a real d-closed 1-form σ on S such that ϕ, $\bar{\phi}$ and σ generate the d-cohomology group of 1-forms on S and such that

$$d\sigma_0 = \phi \wedge \bar{\phi},$$

where σ_0 denotes the component of σ of type $(1,0)$. Moreover there exists on the universal covering W of S a pair of holomorphic functions w_1, w_2 satisfying the differential equations

$$\begin{cases} dw_1 = \phi, \\ dw_2 = \sigma_0 + \bar{w}_1 \phi. \end{cases}$$

We identify $\pi_1(S)$ with the covering transformation group G of W over S in a natural manner. For any covering transformation $g \in G$ we have

$$\begin{cases} gw_1 = w_1 + \alpha, \\ gw_2 = w_2 + \bar{\alpha} w_1 + \beta \end{cases}$$

(see (24)). This formula defines a representation $R: g \to R(g)$ of G by affine transformations of the space $\boldsymbol{C}^2$ of the complex variables w_1 and w_2. Letting

$$g_j w_1 = w_1 + \alpha_j, \quad g_j w_2 = w_2 + \bar{\alpha}_j w_1 + \beta_j, \qquad \text{for } j = 1, 2, 3, 4,$$

and using (67) we obtain

$$\alpha_1 = \alpha_2 = 0, \qquad \bar{\alpha}_3 \alpha_4 - \bar{\alpha}_4 \alpha_3 = m\beta_2.$$

It follows that $m\beta_2 \neq 0$, since the periods α_3, α_4 of the elliptic curve Δ are linearly independent over $\boldsymbol{R}$. Hence the restriction to C of dw_2 is a non-vanishing holomorphic 1-form and its periods β_1, β_2 are linearly independent over $\boldsymbol{R}$. Consequently the representation $R: G \rightarrow R(G)$ is an isomorphism. $R(G)$ is a properly discontinuous group of affine transformations without fixed point of $\boldsymbol{C}^2$ and the quotient space $\boldsymbol{C}^2/R(G)$ is a (compact complex analytic) surface. $dw_1 \wedge dw_2$ is a holomorphic 2-form on S which does not vanish identically, while the canonical bundle of S is trivial. Hence $dw_1 \wedge dw_2$ does not vanish at each point of S. Therefore the pair of holomorphic functions w_1 and w_2 defines a locally biholomorphic map μ of W into $\boldsymbol{C}^2$. The map μ induces a locally biholomorphic map of $S = W/G$ into $\boldsymbol{C}^2/R(G)$ which is surjective in view of the compactness of S. Hence μ is surjective and biholomorphic. Consequently S is complex analytically homeomorphic to $\boldsymbol{C}^2/R(G)$. Thus we obtain the following

THEOREM 19. *Let S be a surface. If the canonical bundle of S is trivial, then S is a* K3 *surface, a complex torus or an elliptic surface of the form $\boldsymbol{C}^2/G$, where $\boldsymbol{C}^2$ is the space of two complex variables (z_1, z_2) and G is a properly discontinuous group of affine transformations without fixed points of $\boldsymbol{C}^2$ which leave invariant the 2-form $dz_1 \wedge dz_2$. The first homology group of the elliptic surface $\boldsymbol{C}^2/G$ is*

$$H_1(\boldsymbol{C}^2/G, \boldsymbol{Z}) \cong \boldsymbol{Z} \oplus \boldsymbol{Z} \oplus \boldsymbol{Z} \oplus \boldsymbol{Z}_m,$$

where $\boldsymbol{Z}_m$ denotes a cyclic group of a finite order m.

THEOREM 20. *Assume that the first Betti number b_1 of S is even and the geometric genus p_g of S is positive. If the first Chern class c_1 of S vanishes, then S is either a* K3 *surface or a complex torus.*

Proof. In case S has no meromorphic function except constants, S is, by Theorem 11, either a K3 surface or a complex torus, since the vanishing of c_1 implies that S contains no exceptional curve. We may assume therefore that S is either algebraic or elliptic. Since p_g is positive, S has an effective canonical divisor K. In view of the above Theorem 19, it suffices for our purpose to prove that K is zero. Since c_1 vanishes, K is homologous to zero. To prove that K is zero, it suffices therefore to show that every positive divisor on S is not homologous to zero. This is obvious if S is algebraic. Hence it suffices to consider the case in which S is a non-algebraic elliptic surface.

Let Δ denote the base curve of the elliptic surface S and let $\{E_j\}$ be a finite covering of Δ by small disks E_j. Since b_1 is even, there exists a complex analytic family of analytic fibre spaces S_t, $t \in C$, of elliptic curves over Δ which has the following three properties:

i) S_0 coincides with S;

ii) S_1 is algebraic;

iii) the complex structure of each piece $S_t \mid E_j$ is independent of t (see [9], § 14).

Consider a positive divisor $D = \sum m_\nu C_\nu$, $m_\nu > 0$, composed of irreducible curves C_ν on S. Each irreducible curve C_ν is a component of a fibre of S (see [8], Theorem 4.3). Hence C_ν is contained in a piece $S \mid E_j$, while $S_t \mid E_j$ is independent of t. Therefore $D = \sum m_\nu C_\nu$ represents a positive divisor on each surface S_t. In particular D is a positive divisor on the algebraic surface S_1. Consequently D is not homologous to zero.

7. **Classification of surfaces.** We shall say that a surface S_1 is *birationally equivalent* to S if there exist quadric transformations $Q_{11}, Q_{12}, \cdots, Q_{1r}, Q_1, Q_2, \cdots, Q_s$ such that

$$Q_{1r} \cdots Q_{12} Q_{11}(S_1) = Q_s \cdots Q_2 Q_1(S).$$

By the *birational class* of S we shall mean the class consisting of all surfaces which are birationally equivalent to S. We shall say that S_1 *dominates* S and write $S_1 \succ S$ if

$$S_1 = Q_s \cdots Q_2 Q_1(S).$$

Clearly the birational class of S forms a partially ordered set with respect to the domination $\succ$. We call S a *relatively minimal surface* if S dominates no surface other than S. It is obvious that S is relatively minimal if and only if S contains no exceptional curve (of the first kind). We call S a *minimal surface* if every surface belonging to the birational class of S dominates S. For each positive integer m we define the *m-genus* $P_m = P_m(S)$ of S to be the dimension of the pluricanonical system $\mid mK \mid$ increased by 1, where K denotes the canonical bundle of S. Note that P_1 coincides with the geometric genus p_g. In case a plurigenus P_m of S is positive, every exceptional curve on S is a fixed component of the pluricanonical system $\mid mK \mid$. Hence we infer that, if S contains no exceptional curve and if a plurigenus P_m of S is positive, then S is a minimal surface.

THEOREM 21. *Surfaces free from exceptional curves can be classified into the following seven classes* (see Table I):

- I_0) *the class of relatively minimal algebraic surfaces with* $p_g=0$;
- II_0) *the class of* K3 *surfaces*;
- III_0) *the class of complex tori* (*of complex dimension* 2);
- IV_0) *the class of minimal elliptic surfaces with* $b_1\equiv 0(2)$, $p_g\geqq 1$, $c_1^2=0$, $c_1\neq 0$;
- V_0) *the class of minimal algebraic surfaces with* $p_g\geqq 1$, $c_1^2>0$;
- VI_0) *the class of minimal elliptic surfaces with* $b_1\equiv 1(2)$, $p_g\geqq 1$, $c_1^2=0$;
- VII_0) *the class of minimal surfaces with* $b_1=q=1$, $p_g=0$.

TABLE I.

class	b_1	p_g	c_1	c_1^2	structure
I_0	even	0			algebraic
II_0	0	+	$=0$	0	K3 surfaces
III_0	4	+	$=0$	0	complex tori
IV_0	even	+	$\neq 0$	0	elliptic
V_0	even	+		+	algebraic
VI_0	odd	+		0	elliptic
VII_0	1	0			?

To prove this theorem we assume S to be a surface free from exceptional curve and denote by p_g, P_m, q, b_1, c_1, $\cdots$ the geometric genus, the m-genus, the irregularity, the first Betti number, the first Chern class, $\cdots$ of S. Moreover we denote by K either the canonical bundle of S or a canonical divisor on S. We remark that, *if an irreducible curve C on S has $(C^2)<0$, then* (KC) is non-negative. In fact, since

$$(KC)+(C^2)=2\pi(C)-2,$$

if (KC) and (C^2) were both negative, then C would be an exceptional curve.

LEMMA 5. *If p_g is positive, then c_1^2 is non-negative.*

Proof. Since, by hypothesis, p_g is positive, there exists on S an effective canonical divisor K. Assuming that $K\neq 0$, we let $K=\sum k_\nu C_\nu$, $k_\nu\geqq 1$, where the C_ν are irreducible curves. We have

$$(KC_\nu)=k_\nu(C_\nu^2)+\sum_{\lambda\neq\nu}k_\lambda(C_\lambda C_\nu)\geqq k_\nu(C_\nu^2).$$

Combining this with the above remark we infer that (KC_ν) is non-negative and, consequently, $c_1{}^2 = \sum k_\nu(KC_\nu)$ is non-negative.

LEMMA 6. *Assume that* $c_1{}^2 = 0$. *Let* $D = \sum k_\nu C_\nu$ *be a positive divisor, where the* C_ν *denote irreducible curves on* S *and the coefficients* k_ν *are positive integers. If* D *belongs to a pluricanonical system* $|mK|$, *then each irreducible component* C_ν *satisfies*

$$(KC_\nu) = 0, \qquad 2\pi(C_\nu) - 2 = (C_\nu{}^2) \leqq 0. \tag{68}$$

Proof. We have

$$m(KC_\nu) = k_\nu(C_\nu{}^2) + \sum_{\lambda \neq \nu} k_\lambda(C_\lambda C_\nu) \geqq k_\nu(C_\nu{}^2).$$

Combining this with the above remark we infer that (KC_ν) is non-negative, while

$$\sum k_\nu(KC_\nu) = mc_1{}^2 = 0.$$

Hence (KC_ν) vanishes and therefore $(C_\nu{}^2)$ is non-positive.

LEMMA 7. *If* $c_1{}^2 = 0$ *and if* $P_m \geqq 2$ *for a positive integer* m, *then* S *is an elliptic surface.*

Proof. By hypothesis the dimension of the pluri-canonical system $|mK|$ is positive. Let A denote the fixed component of $|mK|$. Then any general member of $|mK|$ has the form $A + \sum_\nu C_\nu$, where the C_ν are irreducible curves. Note that $\sum_\nu C_\nu$ is a member of the complete linear system $|mK - A|$ of dimension $\geqq 1$. Consider a general member $\sum_\mu C'_\mu$, $C'_\mu \neq C_\nu$, of $|mK - A|$. By (68), we have

$$\sum_\mu (C'_\mu C_\nu) + (AC_\nu) = m(KC_\nu) = 0$$

and, consequently,

$$(C'_\mu C_\nu) = (AC_\nu) = 0. \tag{69}$$

This shows that the curve $\sum C'_\mu$ does not meet $\sum C_\nu$. Thus any two general members of $|mK - A|$ do not intersect. Consequently the complete linear system $|mK - A|$ is composed of a pencil of curves in the following manner: There is a holomorphic map Ψ of S onto a non-singular algebraic curve Δ such that the inverse image $\Psi^{-1}(u)$ of any general point $u \in \Delta$ is an irreducible non-singular curve and $|mK - A|$ is induced from a complete linear system $|\mathfrak{d}|$ on Δ by means of Ψ. Thus, if $\sum_\nu u_\nu$ is a general member of $|\mathfrak{d}|$, then

$\sum_\nu C_\nu$, $C_\nu = \Psi^{-1}(u_\nu)$, is a general member of $|mK - A|$. Since C_ν does not meet C_λ, $\lambda \neq \nu$, we obtain, using (68) and (69), the equality

$$2\pi(C_\nu) - 2 = (C_\nu^2) = m(KC_\nu) - (AC_\nu) = 0.$$

This proves that $C_\nu = \Psi^{-1}(u_\nu)$ is an elliptic curve. Thus we infer that S is an elliptic surface, q. e. d.

We shall examine various cases separately.

1) If $b_1 \equiv 0 \, (2)$ and $p_g = 0$, then, by Theorem 10, S is an algebraic surface.

2) If $b_1 \equiv 0 \, (2)$, $p_g > 0$ and $c_1 = 0$, then, by Theorem 20, S is either a K3 surface or a complex torus.

3) If $p_g \geqq 2$ and $c_1^3 = 0$, then, by Lemma 7, S is an elliptic surface.

4) If $b_1 \equiv 0 \, (2)$, $p_g = 1$, $c_1^2 = 0$ and $c_1 \neq 0$, then by (13), $8q + b^- = 19$, and therefore q is not greater than 2. Moreover there exists on S a positive canonical divisor $K = \sum k_\nu C_\nu$, $k_\nu \geqq 1$, where the C_ν are irreducible curves.

α) If $q = 0$, then, using (4), we get

$$P_2 = \dim |2K| + 1 \geqq -\dim |-K| + 1 = 2.$$

Hence, by Lemma 7, S is an elliptic surface.

β) If $q = 1$, then the Albanese variety $\mathfrak{A}$ attached to S is an elliptic curve and there is a holomorphic map Φ of S onto $\mathfrak{A}$. Take a general point u of $\mathfrak{A}$ and let $\Gamma_u = \Phi^{-1}(u)$. We infer readily that Γ_u is a non-singular irreducible curve.

If $(K\Gamma_u)$ vanishes, then $2\pi(\Gamma_u) - 2 = (\Gamma_u^2) = 0$ and, consequently, S is an elliptic surface.

If $(K\Gamma_u)$ does not vanish, then one of the irreducible components, say $C = C_1$ of K meets Γ_u. Hence Φ maps C onto $\mathfrak{A}$ and therefore the genus of C is not smaller than 1, while, by (68), the virtual genus $\pi(C)$ is not greater than 1. Consequently C is a non-singular elliptic curve and

$$(KC) = (C^2) = 0. \tag{70}$$

For any complex line bundle F over S we denote by F_C the restriction of F to the curve C. It follows from (70) that C does not meet the curves C_ν, $\nu \geqq 2$. Hence we obtain

$$[K]_C = k[C]_C, \qquad k = k_1,$$

while the canonical bundle $[K+C]_C$ of C is trivial. Setting $m=2k+2$, we infer therefore that $[mK]_C$ is trivial. Hence we have the exact sequence

$$0\rightarrow H^0(S,\mathcal{O}(mK-C))\rightarrow H^0(S,\mathcal{O}(mK))\rightarrow \mathbf{C}\rightarrow H^1(S,\mathcal{O}(mK-C))\rightarrow\cdots,$$

while, by the Riemann-Roch-Hirzebruch theorem,

$$\dim H^0(S,\mathcal{O}(mK-C))-\dim H^1(S,\mathcal{O}(mK-C))=1.$$

Hence we obtain

$$P_m=\dim H^0(S,\mathcal{O}(mK))\geqq 2.$$

Consequently, by Lemma 7, S is an elliptic surface.

γ) If $q=2$, then the Albanese variety $\mathfrak{A}$ attached to S is a complex torus of complex dimension 2 and there exist two linearly independent holomorphic 1-forms ϕ_1, ϕ_2 which determine a holomorphic map Φ of S into $\mathfrak{A}$ (see (15)).

If $\phi_1\wedge\phi_2$ does not vanish identically, then K coincides with the divisor of $\phi_1\wedge\phi_2$:

$$(\phi_1\wedge\phi_2)=K=\sum k_\nu C_\nu.$$

If the image $\Phi(C_\nu)$ of each curve C_ν is a point, then we infer in the same manner as in the proof of Theorem 5.3 in our previous paper [8] that Φ maps S biholomorphically onto $\mathfrak{A}$. This contradicts the assumption that $c_1\neq 0$. Thus we see that at least one of the images, say $\Phi(C_1)$ is a curve in $\mathfrak{A}$. Since, by (68), the virtual genus $\pi(C_1)$ is not greater than 1, C_1 and $\Phi(C_1)$ are both non-singular elliptic curves. There exists therefore a holomorphic map of $\mathfrak{A}$ onto a non-singular elliptic curve which maps $\Phi(C_1)$ onto a point. Hence we infer the existence of a holomorphic map Ψ of S onto a non-singular elliptic curve Δ such that $\Psi(C_1)$ is a point and such that the inverse image $\Psi^{-1}(u)$ of any general point $u\in\Delta$ is a non-singular irreducible curve.

Let $a=\Psi(C_1)$. Since, by (68), $(KC_1)=(C_1^2)=0$, the curve C_1 does not meet C_ν, $\nu\geqq 2$. Moreover *the inverse image* $\Psi^{-1}(a)$ *of* a *coincides with* C_1. To prove this we choose ϕ_1 and ϕ_2 such that ϕ_2 is (induced from) a holomorphic 1-form on Δ and let τ denote a local uniformization variable on Δ such that $d\tau=\phi_2$. Let ζ be an arbitrary point on C_1. Since the restriction of ϕ_1 to C_1 is a non-vanishing holomorphic 1-form, we find a local coordinate (z_1,z_2) of the center ζ on S such that $dz_1=\phi_1$ and z_2 vanishes on C_1. In a neighborhood of ζ, we have

$$(\phi_1\wedge\phi_2)=kC_1,\qquad k=k_1.$$

This means that

$$dz_1 \wedge d\tau(\Psi(z_1, z_2)) = z_2^k f\, dz_1 \wedge dz_2,$$

where f is a non-vanishing holomorphic function on a neighborhood of ζ. It follows that

$$\tau(\Psi(z_1, z_2)) = z_2^{k+1} g + \text{constant},$$

where g is a non-vanishing holomorphic function on a neighborhood of ζ. Therefore, in a neighborhood of ζ, the curve $\Psi^{-1}(a)$ coincides with C_1. This proves that C_1 is a connected component of $\Psi^{-1}(a)$, while, since $\Psi^{-1}(u)$ is irreducible for any general point u, $\Psi^{-1}(a)$ is connected. Hence $\Psi^{-1}(a)$ coincides with C_1.

Since $C_1 = \Psi^{-1}(a)$ does not meet C_ν, $\nu \geqq 2$, the curve $C_u = \Psi^{-1}(u)$ does not meet C_ν for any general point $u \in \Delta$ and therefore

$$2\pi(C_u) - 2 = (KC_u) + (C_u^2) = 0.$$

Consequently S is an elliptic surface.

If $\phi_1 \wedge \phi_2$ vanishes identically, then $\Phi(S)$ is a curve in $\mathfrak{A}$. Hence we infer that there exists a holomorphic map Ψ of S onto a non-singular algebraic curve of genus 2 such that the inverse image $C_u = \Psi^{-1}(u)$ of any general point $u \in \Delta$ is a non-singular irreducible curve. Ψ maps each curve C_ν onto a point on Δ, since, by (68), the genus of C_ν is not greater than 1. Hence C_u does not meet C_ν and therefore

$$2\pi(C_u) - 2 = (KC_u) + (C_u^2) = 0.$$

Consequently S is an elliptic surface.

5) If $c_1^2 > 0$, then, by Theorem 9, S is an algebraic surface.

6) If $b_1 \equiv 1(2)$ and $p_g > 0$, then, by Theorems 4 and 11, S is an elliptic surface.

7) If $b_1 \equiv 1(2)$ and $p_g = 0$, then S is either an elliptic surface or a surface with no meromorphic function (except constants). In case S is elliptic, S has no singular fibre other than that of type ${}_mI_0$, since b_1 is odd (see [9], §11 and §14). Hence

$$12(p_g - q + 1) = c_2 = 0$$

and consequently, $q = b_1 = 1$. In case S has no meromorphic function, we infer from Theorem 11 that $q = b_1 = 1$. We infer readily that S is minimal.

This completes the proof of Theorem 21.

We shall examine the asymptotic behavior of the plurigenera of surfaces with positive geometric genera. We shall indicate by writing $P_m \sim f(m)$ that

$$\lim_{m\to+\infty} P_m/f(m) = 1.$$

Assume S to be a minimal surface with a positive geometric genus of which the canonical bundle is non-trivial. By the above Theorem 21, S is either an elliptic surface or an algebraic surface with $c_1{}^2 > 0$.

In case S is an elliptic surface, the canonical bundle K of S is given by the formula (40), namely,

$$K = \Psi^*(\mathfrak{k} - \mathfrak{f}) + \sum (m_\nu - 1)[P_\nu].$$

Since, by hypothesis, $\dim |\mathfrak{k} - \mathfrak{f}| = p_g - 1 \geqq 0$, the complete linear system $|\mathfrak{k} - \mathfrak{f}|$ on the base curve Δ of S contains an effective divisor $\mathfrak{d}$. Let d denote the degree of $\mathfrak{d}$. For an arbitrary positive integer m, we determine non-negative integers n_ν and r_ν such that

$$m(m_\nu - 1) = n_\nu m_\nu + r_\nu, \qquad r_\nu \leqq m_\nu - 1,$$

and define

$$\mathfrak{d}_m = m\mathfrak{d} + \sum_\nu n_\nu a_\nu, \qquad a_\nu = \Psi(P_\nu).$$

This divisor $\mathfrak{d}_m$ determines a complex line bundle $[\mathfrak{d}_m]$ over Δ and

$$mK = \Psi^*([\mathfrak{d}_m]) + \sum_\nu r_\nu [P_\nu].$$

Hence we obtain

$$P_m = \dim |\mathfrak{d}_m| + 1.$$

Let

$$\kappa = d + \sum_\nu (1 - 1/m_\nu).$$

Since, by hypothesis, K is non-trivial, κ is positive. Therefore, denoting by d_m the degree of $\mathfrak{d}_m$, we get

$$d_m = md + \sum_\nu n_\nu \sim \kappa m.$$

Consequently we obtain

$$P_m \sim \kappa m, \qquad \kappa > 0. \tag{71}$$

In case S is an algebraic surface with $c_1{}^2 > 0$, a theorem of Mumford [16] asserts that the superabundance $\dim H^1(S, \mathcal{O}(mK))$ vanishes for all sufficiently large values of m. Hence, using the Riemann-Roch-Hirzebruch theorem, we obtain

$$P_m = \tfrac{1}{2}m(m-1)c_1^2 + p_g - q + 1 \tag{72}$$

for all sufficiently large values of m and

$$P_m \sim \kappa m^2, \qquad \kappa = \tfrac{1}{2}c_1^2 > 0. \tag{73}$$

We define classes $I, II, \cdots, VII$ *to be respectively the classes of those surfaces which are birationally equivalent to surfaces belonging to the classes* $I_0, II_0, \cdots, VII_0$. Since the numerical characters b_1 and P_m are invariant under quadric transformations, we infer from the above results the following

THEOREM 22. *Surfaces can be classified into seven classes* $I, II, \cdots, VII$. *The surfaces of the classes* $I, II, \cdots, VII$ *are characterized respectively by the following conditions:*

(I) $b_1 \equiv 0(2)$ *and* $p_g = 0$;
(II) $b_1 = 0$, $p_g = 1$ *and* $P_m \sim 1$;
(III) $b_1 = 4$, $p_g = 1$ *and* $P_m \sim 1$;
(IV) $b_1 \equiv 0(2)$, $p_g > 0$ *and* $P_m \sim \kappa m$, $\kappa > 0$;
(V) $b_1 \equiv 0(2)$, $p_g > 0$ *and* $P_m \sim \kappa m^2$, $\kappa > 0$;
(VI) $b_1 \equiv 1(2)$ *and* $p_g > 0$;
(VII) $b_1 = 1$ *and* $p_g = 0$.

THEOREM 23. *Any deformation of a minimal surface with a positive geometric genus is minimal.*

Proof. Since exceptional curves on surfaces are stable submanifolds (see Kodaira [12], Theorem 5), it suffices to show that any sufficiently small deformation of a minimal surface with a positive geometric genus is minimal. Consider a complex analytic family of surfaces S_t, $|t| < 1$, where t denotes several complex variables $(t_1, t_2, \cdots, t_r)$, and assume that S_0 is a minimal surface with $p_g(S_0) > 0$. Note that the numerical characters $p_g = p_g(S_t)$, $q = q(S_t)$ and $c_1^2 = c_1^2(S_t)$ are independent of t. Each surface S_t is obtained from a minimal surface S_t^* by means of a finite number of quadric transformations. Denoting by n_t the number of quadric transformations required to transform S_t^* into S_t, we have

$$c_1^2(S_t^*) = n_t + c_1^2.$$

Since, by the Riemann-Roch inequality (4),

$$P_m(S_t^*) \geqq \tfrac{1}{2}m(m-1)c_1^2(S_t^*) + p_g - q + 1, \qquad \text{for } m \geqq 2,$$

we get

$$P_m(S_t) \geqq \tfrac{1}{2}m(m-1)(n_t+c_1^2)+p_g-q+1, \qquad \text{for } m \geqq 2,$$

while $P_m(S_t)$ is an upper semi-continuous function of t. Hence, for each integer $m \geqq 2$, we find a positive number ϵ_m such that

$$P_m(S_0) \geqq \tfrac{1}{2}m(m-1)(n_t+c_1^2)+p_g-q+1, \qquad \text{for } |t| < \epsilon_m.$$

On the other hand, we infer from (71) and (72) that

$$P_m(S_0) \leqq \tfrac{1}{2}m(m-1)c_1^2+km,$$

where k is a positive constant. Hence we obtain

$$\tfrac{1}{2}m(m-1)n_t+p_g-q \leqq km, \qquad \text{for } |t| < \epsilon_m.$$

This proves that n_t vanishes for $|t| < \epsilon_m$, provided that m is sufficiently large. Consequently the surface S_t is minimal if $|t|$ is sufficiently small, q. e. d.

THEOREM 24. *The classes $II_0, III_0, \cdots, VI_0$ of surfaces are closed under deformations. The classes $I, II, \cdots, VII$ of surfaces are closed under deformations.*

Proof. The classes $I, VI, VII, II_0, III_0, \cdots, VI_0$ are determined respectively by conditions which are invariant under deformations. Hence they are closed under deformations. Considering a complex analytic family of surfaces S_t, $|t| < 1$, with $p_g(S_t) > 0$, we remark that, in view of Theorem 23, the stability of exceptional curves implies the existence of a complex analytic family of minimal surfaces S_t^*, $|t| < \epsilon$, such that

$$S_t = Q_n \cdots Q_2 Q_1(S_t^*),$$

where ϵ is a positive number and $Q_1, Q_2, \cdots Q_n$ denote quadric transformations. This remark shows that the classes $II, III, \cdots, VI$ are also closed under deformations.

THEOREM 25. *A surface is a deformation of an algebraic surface if and only if its first Betti number is even.*

Proof. If S is non-algebraic and if the first Betti number of S is even, then, by Theorem 22, either S is an elliptic surface or S is obtained from a complex torus or a K3 surface by means of a finite number of quadric transformations. An elliptic surface is a deformation of an algebraic surface if its first Betti number is even (see [9], Theorem 15.2). Any complex torus is obviously a deformation of an algebraic surface. By Theorem 13 every K3 surface is a deformation of a non-singular quartic surface in a projective 3-space. Consequently S is a deformation of an algebraic surface.

REFERENCES.

[1] M. F. Atiyah and F. Hirzebruch, "Riemann-Roch theorems for differentiable manifolds," *Bulletin of the American Mathematical Society*, vol. 65 (1959), pp. 276-281.

[2] M. F. Atiyah and I. M. Singer, "The index of elliptic operators on compact manifolds," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 422-433.

[3] W. L. Chow and K. Kodaira, "On analytic surfaces with two independent meromorphic functions," *Proceedings of the National Academy of Sciences of the U.S.A.*, vol. 38 (1952), pp. 319-325.

[4] P. Dolbeault, "Sur la cohomologie des variétés analytiques complexes," *Comptes Rendus des Séances de l'Académie des Sciences*, vol. 236 (1953), pp. 175-177.

[5] H. Grauert, "On the number of moduli of complex structures," in *Contributions to Function Theory*, Tata Institute, 1960.

[6] ———, "Über Modifikationen und exzeptionelle analytische Mengen," *Mathematische Annalen*, vol. 146 (1962), pp. 331-368.

[7] F. Hirzebruch, *Neue topologische Methoden in der algebraischen Geometrie*, zweite Auflage, Springer, Berlin, 1962.

[8] K. Kodaira, "On compact complex analytic surfaces, I," *Annals of Mathematics*, vol. 71 (1960), pp. 111-152.

[9] ———, "On compact analytic surfaces, II-III," *Annals of Mathematics*, vol. 77 (1963), pp. 563-626, vol. 78 (1963), pp. 1-40.

[10] ———, "On the structure of compact complex analytic surfaces," *Proceedings of the National Academy of Sciences of the U.S.A.*, vol. 50 (1963), pp. 218 221.

[11] ———, "On the structure of compact complex analytic surfaces, II," *Proceedings of the National Academy of Sciences of the U.S.A.*, vol. 51 (1964), pp. 1100-1104.

[12] ———, "On stability of compact submanifolds of complex manifolds," *American Journal of Mathematics*, vol. 85 (1963), pp. 79-94.

[13] ———, L. Nirenberg and D. C. Spencer, "On the existence of deformations of complex analytic structures," *Annals of Mathematics*, vol. 68 (1958), pp. 450-459.

[14] ——— and D. C. Spencer, "On deformations of complex analytic structures, I, II, III,' *Annals of Mathematics*, vol. 67 (1958), pp. 328-466, vol. 71 (1960), pp. 43-76.

[15] ——— and ———, "A theorem of completeness for complex analytic fibre spaces," *Acta Mathematica*, vol. 100 (1958), pp. 281-294.

[16] D. Mumford, "The canonical ring of an algebraic surface," *Annals of Mathematics*, vol. 76 (1962), pp. 612-615.

[17] Th. Skolem, *Diophantische Gleichungen*, Springer, Berlin, 1938.

ON UNMIXEDNESS THEOREM.*

By Wei-Liang Chow.

B. Dwork and K. Ireland, in their recent work on the zeta function of an algebraic variety, have raised the question whether the unmixedness theorem holds for the doubly or multiply projective space, or rather for the (generalized) Segre representation of such a space as a projective variety. In this paper we shall give an affirmative answer to this question. However, in course of our investigation, it turns out that our method is susceptible of a bit of generalization; in fact, we shall now express our result as a theorem on the validity of the unmixedness theorem for the "Segre" product of two Noetherian graded rings, assuming that unmixedness theorem holds for the two rings themselves. This generalization is not nominal, for, apart from giving a clearer picture of the nature of our reasoning, it yields also some additional results of quite a bit of interest, as we shall see presently.

1. Let V be a projective variety, defined over a field K, and let R be the homogeneous coordinate ring of V. We shall say that the unmixedness theorem holds for V, if the unmixedness theorem holds for homogeneous ideals in R. We recall that the unmixedness theorem is said to hold for homogeneous ideals if the following assertion is true: If a homogeneous ideal $\mathfrak{a}$ in R generated by r elements has the height r, then it is unmixed, i.e., has no embedded prime divisor. Let $\mathfrak{M}$ be the (maximal) ideal in R generated by all elements of positive degrees, and let $R_{\mathfrak{M}}$ be the quotient ring of R with respect to $\mathfrak{M}$; since every prime divisor of $\mathfrak{a}$ is homogeneous and hence must be contained in $\mathfrak{M}$, it follows that $\mathfrak{a}$ is unmixed if and only if $R_{\mathfrak{M}}\mathfrak{a}$ is unmixed. Now, it is easily seen that $R_{\mathfrak{M}}$ has a system of parameters consisting of homogeneous elements in R; hence the validity of the unmixedness theorem for homogeenous ideals implies that any such system of parameters must be distinct. Conversely, it is well-known that if $R_{\mathfrak{M}}$ possesses one distinct system of parameters (in which case $R_{\mathfrak{M}}$ is said to be Macaulay), then every system of parameters in $R_{\mathfrak{M}}$ must be distinct and the unmixedness theorem holds in $R_{\mathfrak{M}}$. Thus we see that the unmixedness theorem holds for homogeneous ideals in R if and only if $R_{\mathfrak{M}}$ is Macaulay; in other words, the unmixedness theorem holds for V if and only if $R_{\mathfrak{M}}$ is Macaulay.

---

* This work was partially supported by the National Science Foundation under Grant NSF-G7030.

Now, let U be another projective variety, defined over the same field K; let S be the homogeneous coordinate ring of U, and let $\mathfrak{N}$ be the ideal in S generated by all elements of positive degrees. Let T be the tensor product of R and S over K; we note that both R and S can be canonically embedded as subrings in T, and that their gradings induce in T a bi-grading, so that T is a bigraded ring. If we set $R=\sum_{i=0}^{\infty} R_i$ and $S=\sum_{i=0}^{\infty} S_i$, where R_i and S_i are the K-modules of homogeneous elements of degree i in R and S respectively, then we have $T=\sum_{i,j=0}^{\infty} T_{i,j}$ with $T_{i,j}=R_iS_j$, and the ideal $\mathfrak{D}=T(\mathfrak{M},\mathfrak{N})$ is the maximal ideal in T generated by all homogeneous elements of positive bidegrees. As we shall see later, it can be easily shown that if both $R_{\mathfrak{M}}$ and $S_{\mathfrak{N}}$ are Macaulay, then $T_{\mathfrak{D}}$ is also Macaulay. Now, T can be considered as the homogeneous coordinate ring of a projective variety $J(V,U)$ which is sometimes called the join of V and U; briefly, $J(V,U)$ may be described as the variety generated by all the lines joining a point of V with a point of U, assuming that the ambient projective spaces of V and U are embedded as two skew linear subspaces in a third (bigger) projective space. Thus, if the unmixedness theorem holds for both V and U, then it holds also for $J(V,U)$. However, we are interested here not so much in the join $J(V,U)$ as in the product variety $V\times U$ represented projectively as a subvariety in the Segre variety representing the product of their ambient spaces; we shall call the so represented $V\times U$ the Segre product of V and U. More generally, we shall consider the Segre product $V^{(d)}\times U^{(e)}$, where $V^{(d)}$ and $U^{(e)}$ are the images of V and U respectively under birational transformations determined by the linear systems of hypersurfaces of degrees d and e respectively. The homogeneous coordinate ring of $V^{(d)}\times U^{(e)}$ is the subring $A=\sum_{i=0}^{\infty} T_{id,ie}$ in T generated by all homogeneous elements of bi-degrees (id,ie) for all $i=0,1,\cdots$; and the ideal $\mathfrak{Q}=A\cap\mathfrak{D}$ is the maximal ideal in A generated by all homogeneous elements of positive degrees. Our main problem is to determine the condition under which the ring $A_{\mathfrak{Q}}$ is Macaulay, assuming that both $R_{\mathfrak{M}}$ and $S_{\mathfrak{N}}$ are Macaulay. The conditions we shall give will be easily seen to be satisfied in case both R and S are polynomial rings over K, endowed with their usual grading; and since polynomial rings are homogeneous coordinate rings of projective spaces, this answers the question posed at the beginning of this paper.

However, as mentioned above, our investigations can be extended readily to more general situations. Instead of homogeneous K-algebras arising from algebraic varieties, we shall only assume that R and S are Noetherian graded algebras over a common field K (with the restrictions $R_0=S_0=K$); it is easily seen that everything we have said above, except for the geometrical interpretations, holds without any change in this more general situation, provided we assume that $R_0=S_0=K$. The main point in this generalization is that the rings involved are no longer assumed to be generated by the homogeneous elements of the first degree. This last point enables us to apply our result to the following case, which I understand from Dwork is of interest in his study of the zeta-function of a complete intersection: We take for R and S as before the polynomial rings over K, but for one (or both) of these two rings we shall assign a different grading from the usual one; for example, if $y_1, \cdots, y_n$ are the variable generating ring S, we shall assign to each y_i an arbitrary positive integer g_i as its degree. It is clear that this case is a generalization of the case of two projective spaces, and we shall see later that our main result remains valid here.

We shall follow generally the terminology of Nagata in his recent book on *Local Rings* (referred to as LR) since this terminology represents what is now more or less standard in the field of commutative algebra. The only exception is that we shall speak of a homogeneous ideal in a graded ring instead of a graded ideal, which is the term used in LR. For the sake of convenience of the reader, we shall give occasional reference to LR for some of the well-known results used in our argument. In this connection we should like to point out that, in course of our investigation, we have found it convenient to introduce the concepts of an H-local ring and of an H-Macaulay ring. These are natural generalizations to the case of a graded ring of the usual concepts of a local ring and of a Macaulay ring, and they are more convenient to use in questions concerning homogeneous elements and ideals in a graded ring. We shall develop these notions here only to the extent needed for our present purpose, but it is likely that they may prove also useful in other connections.

2. In this section we shall develop some results about a certain type of ideals in a Noetherian ring, which we shall need later. Let R be a Notherian ring; we recall that a sequence of elements $u_1, \cdots, u_r$ in R is said to be an R-sequence if $R(u_1, \cdots, u_r) \neq R$ and if u_i is not a zero-divisor for the R-module $R/R(u_1, \cdots, u_{i-1})$ for every $i=1, \cdots, r$, and that this property is independent of the arrangement of the elements of the system into a sequence, so that we can speak of a (finite) set of elements forming an R-sequence without any specification about their arrangement into a sequence.

We begin with the simple observation that if two elements u and v in R form an R-sequence, then $R(uv) = R(u) \cap R(v)$. In fact, if an element au in $R(u)$ is contained in $R(v)$, then au corresponds to the zero element in the R-module $R/R(v)$; since u is not a zero-divisor for $R/R(v)$, a itself must correspond to the zero element and hence is contained in $R(v)$; this shows that au is contained in $R(uv)$. We are interested in generalizing this simple result to a situation which involves two systems of edements as well as more intricate conditions. Let $u_1, \cdots, u_t$ and $v_1, \cdots, v_t$ be two systems of elements in R satisfying the following two conditions, where m and n are two fixed positive integers $\leqq t$: (A) Every set containing at most m elements among the $u_1, \cdots, u_t$ and at most n elements among the $v_1, \cdots, v_t$ forms an R-sequence, and (B) the ideal $R(u_1, \cdots, u_t)$ is generated by any set of m elements among the $u_1, \cdots, u_t$, and the ideal $R(v_1, \cdots, v_t)$ is generated by any set of n elements among the $v_1, \cdots, v_t$. We set $w_i = u_i v_i$, $i = 1, \cdots, t$, and consider the ideal $R(w_1, \cdots, w_t)$; we shall show (Lemma 3) that, in case $t < m + n - 1$, this ideal can be expressed as the intersection of all ideals of the form $R(u_{k_1}, \cdots, u_{k_\alpha}, v_{k_{\alpha+1}}, \cdots, v_{k_t})$, where the indices $k_1, \cdots, k_t$ are a rearrangement of the integers from 1 to t. Before we do this, we shall first prove a result (Lemma 1) which will be useful to us later. In what follows in this section, we shall set for convenience

$$R(u) = R(u_1, \cdots, u_m), \qquad R(v) = R(v_1, \cdots, v_n),$$
$$R(u, v) = R(u_1, \cdots, u_m, v_1, \cdots, v_n),$$

and we observe that by virtue of (B), we have the relatons

$$R(u) = R(u_1, \cdots, u_t), \qquad R(v) = R(v_1, \cdots, v_t),$$
$$R(u, v) = R(u_1, \cdots, u_t, v_1, \cdots, v_t).$$

LEMMA 1. *Asume that $t \leqq m + n - 2$, and let z be an element in R which is contained every one of the ideals $R(u_{k_1}, \cdots, u_{k_\alpha}, v_{k_{\alpha+1}}, \cdots, v_{k_t})$, for all partitions of the integers from 1 to t into two sets $(k_1, \cdots, k_\alpha)$ and $(k_{\alpha+1}, \cdots, k_t)$ and for all values of α from $t - n + 1$ to $m - 1$. Then there exist elements $c_{i_1 \cdots i_m}$ in R such that*

$$z = \sum_{i_1 < \cdots < i_m} c_{i_1 \cdots i_m} v_{i_1} \cdots v_{i_m} \bmod R(u)$$

Proof. Since z is evidently contained in $R(u, v)$, we can set $z = z' + z''$, where z' is contained in $R(u)$ and z'' is contained in $R(v)$; we set $z'' = \sum_{i=1}^{n} c_i v_i$.

Consider a typical term $c_i v_i$, in z'' and take an ideal

$$I_i = R(u_{k_1}, \cdots, u_{k_{m-1}}, v_{k_m}, \cdots, v_{k_t}),$$

where the $t-m+1$ $(\leqq n-1)$ indices $k_m, \cdots, k_t$ are to be taken among the integers from 1 to n but distinct from the index i; we denote by J_i the ideal generated in R by all the elements $v_1, \cdots, v_n$ except v_i. Since z is by assumption contained in I_i, the element $c_i v_i + z' = z - \sum_{j \neq i} c_j v_j$ must be contained in $I_i + J_i$, and hence the element $c_i v_i$ must be contained in $R(u) + J_i$. Since $(R(u) + J_i) : R(v_i) = R(u) + J_i$, it follows that c_i is contained in $R(u) + J_i$; hence, after transferring to z' a term belonging to $R(u)$ if necessary, we can set $z'' = \sum_{i<j} c_{ij} v_i v_j$, where the indices i and j run from 1 to n. If $m = 2$, we can stop here for we will have already arrived at the desired expression for z''; if $m > 2$, we can proceed as follows. Consider a typical term $c_{ij} v_i v_j$ in z'', and take an ideal $I_{ij} = R(u_{k_1}, \cdots, u_{k_{m-1}}, v_{k_m}, \cdots, v_{k_t})$, where the $t-m+1$ indices $k_m, \cdots, k_t$ are now to be taken among the integers from 1 to n or $n+1$ in case $t-m+1 \leqq n-2$ or $t-m+1 > n-2$ respectively, but distinct from both i and j; we observe that since $m > 2$, the condition $t-m+1 > n-2$ implies $t > n$, so that element v_{n+1} exists. We denote by J_{ij} the ideal generated in R by all the elements $v_1, \cdots, v_n$ or $v_1, \cdots, v_{n+1}$ in the respective case except v_i and v_j. Since z is by assumption contained in I_{ij}, the element $c_{ij} v_i v_j + z'$ must be contained in $I_{ij} + J_{ij}$, and hence the element $c_{ij} v_i v_j$ must be contained in $R(u) + J_{ij}$. Since

$$(R(u) + J_{ij}) : R(v_i) = R(u) + J_{ij} \text{ and } (R(u) + J_{ij}) : R(v_j) = R(u) + J_{ij},$$

it follows that c_{ij} is contained in $R(u) + J_{ij}$; hence, after transferring to z' a term belonging to $R(u)$ if necessary, we can set $z'' = \sum_{i<j<k} c_{ijk} v_i v_j v_k$, where the indices run from 1 to n or to $n+1$ in case $t-m+1 \leqq n-2$ or $t-m+1 > n-2$ respectively.

Thus we can proceed step by step. In general, the situation can be described as follows. If $t-m+1 = n-s+1$ $(2 \leqq s \leqq n)$, no new v's need be introduced in addition to $v_1, \cdots, v_n$ during the first s steps; for, as long as the degree of the expression for z'' is less than s, the $t-m+1$ elements $v_{k_m}, \cdots, v_{k_t}$ can be chosen among the $v_1, \cdots, v_n$ and still be distinct from those involved in any one given term in the expression for z''. Thus, after the first s steps (note that $t-m+1 = n-s+1$ implies $s \leqq m$), we obtain for z'' an expression of the form $z'' = \sum_{i_1 < \cdots < i_s} c_{i_1 \cdots i_s} v_{i_1} \cdots v_{i_s}$, where the indices all run from 1 to n. After that, if $s < m$, we can proceed further, but we need to introduce one new element among the $v_{n+1}, \cdots, v$ at each step. In fact, at he beginning of the $(s+1)$-th step there are only $n-s$ elements among the $v_1, \cdots, v_n$ which are distinct from those involved in any one given term in the expression for z''; therefore we need to bring in v_{n+1} to make up the set of $t-m+1=n-s+1$ elements $v_{k_m},\cdots,v_{k_t}$, and we observe that this is possible since $s<m$ implies $t-s+1>t-m+1=n-s+1$ and hence $t>n$. Generally, for any integer r such that $s<r<m$, we have after the r-th step an expression for z'' of the form $z''=\sum_{i_1<\cdots<i_r} c_{i_1\cdots i_r}v_{i_1}\cdots,v_{i_r}$, where the indices run from 1 to $n+r-s$. For the $(r+1)$-th step, since there are only $(n+r-s)-r=n-s$- elements among the $v_1,\cdots,v_{n+r-s}$ which are distinct from those involved in any one given term in the expression for z'', we need to bring in $v_{n+r-s+1}$ to make up the set $v_{k_m},\cdots,v_{k_t}$; and again we observe that this is always possible since $r<m$ implies $t-r+1>n-s+1$ and hence $t>n+r-s$. Thus we can always proceed until after m steps we obtain for z'' an expression of the form $z''=\sum_{i_1<\cdots<i_m} c_{i_1\cdots i_m}v_{i_1}\cdots v_{i_m}$; this proves our lemma.

LEMMA 2. *We have the relation $R(u)R(v)=R(u)\cap R(v)$; in particular, in case either $m=1$ or $n=1$ we have the relation $R(w)=R(u)\cap R(v)$.*

Proof. To prove the relation $R(u)R(v)=R(u)\cap R(v)$, we have to show that $R(u)\cap R(v)$ is the ideal generated by all elements of the form u_iv_j, for $i=1,\cdots,m$, and $j=1,\cdots,n$. It is clearly sufficient to show that if $z=\sum_{j=1}^{n}c_jv_j$ is an element in $R(v)$ which is also contained in $R(u)$, then we can choose the elements c_j so that they all are contained $R(u)$. Assume that for any given integer k $(0<k<n)$ the elements c_j with $j<k$ are all contained in $R(u)$, we shall that c_k can be so chosen (without affecting the c_j with $j<k$) that it is also contained in $R(u)$; this would of course prove our assertion. In fact, since z as well as all c_j with $j<k$ are contained in $R(u)$, the element $c_kv_k=z-\sum_{j\neq k}c_jv_j$ is contained in $R(u)+R(v_{k+1},\cdots,v_n)$; and since $(R(u)+R(v_{k+1},\cdots,v_n)):R(v_k)=R(u)+R(v_{k+1},\cdots,v_n)$, it follows that c_k is contained in $R(u)+R(v_{k+1},\cdots,v_n)$. If we set $c_k=c_k'+\sum_{j>k}c_j'v_j$, where c_k' is contained in $R(u)$, then we have for z the expression $z=\sum_{j<k}c_jv_j+c_k'v_k+\sum_{j>k}(c_j+c_j'v_k)v_j$, in which the coefficients of all v_j with $j<k+1$ are contained in $R(u)$; our assertion is therefore proved. The last statement of the lemma follows by observing that, in case $m=1$, the element u_iv_j is contained in $R(u_jv_j)=R(w_j)$ for every $i=1,\cdots,t$, and hence $R(w)=R(u)R(v)$.

LEMMA 3. *In case $t<m+n-1$, then*

$$(1)\qquad R(w_1,\cdots,w_t)=\bigcap R(u_{i_1},\cdots,u_{i_\alpha},\ v_{i_{\alpha+1}},\cdots,v_{i_t})\cap R(u)\cap R(v),$$

where the intersection extends over all partitions of the integers from 1 *to* t *into two sets* $(i_1, \cdots, i_\alpha)$ *and* $(i_{\alpha+1}, \cdots, i_t)$, *for all values of* α *from* $t-n+1$ *to* $m-1$.

Proof. We use double induction on the pair of integers (m, n); as initial step of our induction, we take the case where either $m=1$ or $n=1$, and observe that in this case the formula (1) reduces to the formula $R(w) = R(u) \cap R(v)$ proved already in Lemma 2. We therefore assume that both m and n are greater than 1, and that our lemma holds true in case either m or n is replaced by a smaller integer. It is sufficient to show that

$$(2) \qquad R(w_1, \cdots, w_t) = R(w_1. \cdots, w_{t-1}, u_t) \cap R(w_1, \cdots, w_{t-1}, v_t).$$

To see this, consider the residue ring $\bar{R} = R/R(v_t)$ and denote generally by a bar the image in $\bar{R}$ of an element in R; it is easily seen that the systems $\bar{u}_1, \cdots, \bar{u}_{t-1}$ and $\bar{v}_1, \cdots, \bar{v}_{t-1}$ satisfy the conditions (A) and (B) with the integers t, m, n replaced either by $t-1, m, n-1$ respectively in case $t > m$, or by $t-1, m-1, n-1$ respectively in case $t = m$. In case $t > m$, we have the inequality $t-1 < m + (n-1) - 1$, while in case $t = m$ the inequality $t-1 < (m-1) + (n-1) - 1$ will also be satisfied except when $n = 2$. Thus, except when $t = m$, $n = 2$, the induction hypothesis is always satisfied; on the other hand, in case $t = m$, $n = 2$, we have the relation

$$\bar{R}(\bar{w}_1, \cdots, \bar{w}_{t-1}) = \bar{R}(\bar{u}) \cap \bar{R}(\bar{v}) = \bar{R}(\bar{u}_1, \cdots, \bar{u}_{m-1}) \cap \bar{R}(\bar{v}_1)$$

by the last assertion of Lemma 2, which is a special case of formula (1), as we have just noted above. Therefore, either by induction hypothesis or by Lemma 2, we have the relation

$$(3) \qquad \begin{aligned} \bar{R}(\bar{w}_1, \cdots, \bar{w}_{t-1}) = {} & \cap \bar{R}(\bar{u}_{i_1}, \cdots, \bar{u}_{i_\alpha}, \bar{v}_{i_{\alpha+1}}, \cdots, \bar{v}_{i_{t-1}}) \\ & \cap \bar{R}(\bar{u}_1, \cdots, \bar{u}_{t-1}) \cap \bar{R}(\bar{v}_1, \cdots, \bar{v}_{t-1}) \end{aligned}$$

where the intersection extends over all partitions of the integers from 1 to $t-1$ into two sets $(i_1, \cdots, i_\alpha)$ and $(i_{\alpha+1}, \cdots, i_{t-1})$, for all values of α from $t-n+1$ to $\inf(m-1, t-2)$. Going over to the ring R, we obtain the relation

$$(4) \qquad \begin{aligned} & R(w_1, \cdots, w_{t-1}, v_t) = \\ & \quad \bigcap R(u_{i_1}, \cdots, u_{i_\alpha}, v_{i_{\alpha+1}}, \cdots, v_{i_{t-1}}, v_t) \cap R(u_1, \cdots, u_{t-1}, v_t) \\ & \quad \cap R(v_1, \cdots, v_t). \end{aligned}$$

Similarly, applying the same argument to the rseidue ring $R/R(u_t)$, we obtain the relation

$$
\begin{aligned}
&R(w_1,\cdots,w_{t-1},u_t) = \\
(5)\qquad &\quad \bigcap R(u_{j_1},\cdots,u_{j_\beta}, v_{j_{\beta+1}},\cdots,v_{j_{t-1}},u_t) \cap R(u_1,\cdots,u_t) \\
&\qquad\qquad \cap R(v_1,\cdots,v_{t-1},u_t),
\end{aligned}
$$

where β runs from $\sup(t-n,1)$ to $m-2$. If we substitute (4) and (5) into (2) and eliminate some redundant terms, we obtain (1).

To prove (2), consider an element z in R such that zu_t is contained in $R(w_1,\cdots,w_{t-1},v_t)$; we have to show that there exists an element z_1 in $R(w_1,\cdots,w_{t-1},v_t)$ such that $zu_t-z_1u_t$ is contained in $R(w_1,\cdots,w_{t-1})$. Going over to the residue ring $\bar R$ introduced above, our assumption implies that $\bar z\bar u_t$ is contained in $\bar R(\bar w_1,\cdots,\bar w_{t-1})$; in view of (3), this implies that $\bar z\bar u_t$ is contained in each $\bar R(\bar u_{i_1},\cdots,\bar u_{i_\alpha}, \bar v_{i_{\alpha+1}},\cdots,\bar v_{i_{t-1}})$ as well as in $\bar R(\bar v_1,\cdots,\bar v_{t-1})$. Since $\bar u_{i_1},\cdots,\bar u_{i_\alpha}, \bar v_{i_{\alpha+1}},\cdots,\bar v_{i_{t-1}},\bar u_t$ forms an $\bar R$-sequence, it follows that $\bar z$ must be contained in $\bar R(\bar u_{i_1},\cdots,\bar u_{i_\alpha}, \bar v_{i_{\alpha+1}},\cdots,\bar v_{i_{t-1}})$; similarly, we conclude that $\bar z$ is contained in $\bar R(\bar v_1,\cdots,\bar v_{t-1})$. If $t=m$, then $\bar u_1,\cdots,\bar u_{t-1}$, $\bar u_t$ also form an $\bar R$-sequence, and hence $\bar z$ is contained also in $\bar R(\bar u_1,\cdots,\bar u_{t-1})$; it follows then from (3) that $\bar z$ is contained in $\bar R(\bar w_1,\cdots,\bar w_{t-1})$, and hence z is contained in $R(w_1,\cdots,w_{t-1},v_t)$, so that we can take z_1 to be z itself. We can therefore assume that $t>m$. According to Lemma 1, applied to the ring $\bar R$, the element $\bar z$ can be expressed in the form $\bar z'+\bar z''$, where $\bar z'$ is contained in $\bar R(u_1,\cdots,\bar u_{t-1})$ and $\bar z''$ has the form

$$\sum_{i_1<\cdots<i_m} \bar c_{i_1\cdots i_m}\bar v_{i_1}\cdots\bar v_{m}.$$

Going over to the ring R, we conclude that $z=z'+z''$, where z' is contained in $R(u_1,\cdots,u_{t-1},v_t)$ and z'' has the form $\sum_{i_1<\cdots<i_m} c_{i_1\cdots i_m}v_{i_1}\cdots v_{i_m}$, where the indices run from 1 to $t-1$. Now, since u_t is contained in $R(u_{i_1},\cdots,u_{i_m})$, for every such set of indices $i_1,\cdots,i_m$, it follows that $v_{i_1}\cdots v_{i_m}u_t$ is contained in $R(w_{i_1},\cdots,w_{i_m})$, and hence $z''u_t$ is contained in $R(w_1,\cdots,w_{t-1})$; this shows that $zu_t-z'u_t$ is contained in $R(w_1,\cdots,w_{t-1})$. Going back again to the residue ring $\bar R$, we see that $\bar z'\bar u_t$ is also contained in $\bar R(w_1,\cdots,w_{t-1})$; hence, by the same argument as we have applied to $\bar z\bar u_t$ before, we conclude that $\bar z'$ is contained in every $\bar R(\bar u_{i_1},\cdots,\bar u_{i_\alpha}, \bar v_{i_{\alpha+1}},\cdots,\bar v_{i_{t-1}})$ and in $\bar R(\bar v_1,\cdots,\bar v_{t-1})$. Since $\bar z'$ is also contained in $\bar R(\bar u_1,\cdots,\bar u_{t-1})$, it follows from (3) that $\bar z'$ is contained in $\bar R(\bar w_1,\cdots,\bar w_{t-1})$. Going over to R, this means that z' is contained in $R(w_1,\cdots,w_{t-1},v_t)$; this concludes the proof of Lemma 3.

Lemma 3 ceases to be generally valid in case $t\geqq m+n-1$; in fact, in this case the right side of (1) reduces to $R(u)\cap R(v)$, which is by Lemma 2 the ideal generated by all the elements of the form u_iv_j, while the left side can be a smaller ideal, as can be seen from the following simple example. Let $m = n = 2$, and $t = 3$; take for u_1, u_2, v_1, v_2 any set of four elements forming an R-sequence, and set $u_3 = u_1 + u_2$, $v_3 = v_1 + v_2$. Then $R(w_1, w_2, w_3)$ is the ideal generated by the elements u_1v_1, u_2v_2, $u_1v_2 + u_2v_1$, which is smaller than the ideal $R(u) \cap R(v) = R(u_1, u_2) \cap R(v_1, v_2) = R(u_1v_1, u_2v_2, u_1v_2, u_2v_1)$. On the other hand, it can be easily verified that every element in $R(u) \cap R(v)$ which can be expressed as a form of degree 3 in the u_1, u_2, v_1, v_2 is contained in $R(w_1, w_2, w_3)$; for example, the element

$$\begin{aligned} u_1^2v_2 &= u_1(u_3 - u_2)v_2 = u_1u_3v_2 - u_1w_2 \\ &= u_1u_3(v_3 - v_1) - u_1w_2 = u_1w_3 - u_3w_1 - u_1w_2 \end{aligned}$$

is contained in $R(w_1, w_2, w_3)$. Therefore, if we denote by $R(w, (u, v)^3)$ the sum of the ideal $R(w_1, w_2, w_3)$ with the ideal $R(u_1, u_2, v_1, v_2)^3$, we obtained the equation $R(w_1, w_2, w_t) = R(u) \cap R(v) \cap R(w, (u, v)^3)$. In case R is the polynomial ring $K[u_1, u_2, v_1, v_2]$ over a field K generated by the indeterminates u_1, u_2, v_1, v_2, this equation gives a primary decomposition of $R(w_1, w_2, w_3)$, in which $R(w, (u, v)^3)$ is an embedded component belonging to the maximal prime ideal $R(u_1, u_2, v_1, v_2)$. The situation we have just described holds in general for $t = m + n - 1$.

LEMMA 4. *For* $t = m + n - 1$, *we have the relation*

$$R(w_1, \cdots, w_t) = R(u) \cap R(v) \cap R(w, (u, v)^t), \tag{6}$$

where $R(w, (u, v)^t) = R(w_1, \cdots, w_t) + R(u, v)^t$.

Proof. It is sufficient to show that every element z in $R(u) \cap R(v)$ which is a "monomial" of degree t is contained in $R(w_1, \cdots, w_t)$. Since $t = m + n - 1$, the monomial z must have either a degree $\geqq n$ in the u's or a degree $\geqq m$ in the v's; assume the first case and set $z = au_{i_1} \cdots u_{i_n}v_j$. We maintain that every such monomial can be expressed as a sum of such monomials in which all the indices $i_1, \cdots, i_n$ are distinct. In fact, assume that $i_1, \cdots, i_\alpha$ $(\alpha < n)$ are distinct and let $k_1, \cdots, k_m$ be m distinct positive integers $\leqq t$ which are distinct from $i_1, \cdots, i_\alpha$ (observe that $t - \alpha \geqq m$), then we can express every one of the factors $u_{i_{\alpha+1}}, \cdots, u_{i_n}$ in z by a linear combination of $u_{k_1}, \cdots, u_{k_m}$, and obtain an expression of z as a sum of monomials each of which contains at least $\alpha + 1$ distinct u's; after a finite number of steps, we arrive at an expression of z as a sum of monomials each of which contains n distinct u's. Thus we only need to prove our assertion for the case where the indices $i_1, \cdots, i_n$ in the expression for z are distinct;

but in this case we can express v_j as a linear combination of $v_{i_1}, \cdots, v_{i_n}$, say $v_j = b_1 v_{i_1} + \cdots + b_n v_{i_n}$, and obtain the equation

$$\begin{aligned} z &= a u_{i_1} \cdots u_{i_n}(b_1 v_{i_1} + \cdots + b_n v_{i_n}) \\ &= a b_1 u_{i_2} \cdots u_{i_n} w_{i_1} + \cdots + a b_n u_{i_1} \cdots u_{i_{n-1}} w_{i_n} \end{aligned}$$

which shows that z is contained in $R(w_1, \cdots, w_t)$.

Remark. Since the above proof uses only the condition (B) but not the condition (A), Lemma 4 therefore holds also without the condition (A).

3. Let R be a Noetherian graded ring, and let $R = \sum_{i=0}^{\infty} R_i$ be the expression of R as a direct sum of modules of homogeneous elements of different degrees; it is clear that R_0 is a Notherian subring in R and that each R_i is also an R_0-module. We shall say that R is *H-local* if it has only one maximal homogeneous ideal. It is easily seen that R is H-local if and only if the subring R_0 is local, and that if such is the case and if $\mathfrak{M}$ is the maximal homogeneous ideal of R, then $\mathfrak{M}_0 = R_0 \cap \mathfrak{M}$ is the maximal ideal of the local ring R_0. The concept of an H-local ring is a natural generalization of that of a local ring (and reduces to it in case $R = R_0$), and as such it has many of the properties of a local ring provided we consider only homogeneous ideals. We shall introduce here some notions and a few results about an H-local ring which we shall need for our present purpose; these results can be proved easily by generalizing the proofs of the corresponding results for a local ring, but it is even easier to derive the former from the latter by using the basic relation $\mathfrak{a} = R \cap R_{\mathfrak{M}}\mathfrak{a}$ for any homogeneous ideal $\mathfrak{a}$ in R. This relation $\mathfrak{a} = R \cap R_{\mathfrak{M}}\mathfrak{a}$ for a homogeneous ideal $\mathfrak{a}$ in R follows from (and is in fact equivalent to) the well-known fact (LR, 8.7) that every homogeneous ideal in R has a primary decomposition consisting of only homogeneous primary ideals, which together with their associated prime ideals must then all be contained in $\mathfrak{M}$. If m is the height of the maximal homogeneous ideal $\mathfrak{M}$ of an H-local ring R, then a system of elements $x_1, \cdots, x_m$ in R is said to be a *system of parameters* for $\mathfrak{M}$ if the ideal $R(x_1, \cdots, x_m)$ is primary for $\mathfrak{M}$; and a system of parameters for $\mathfrak{M}$ is said to be *distinct* if it forms an R-sequence. We shall consider here only system of *homogeneous* parameters for $\mathfrak{M}$, i.e. a system of parameters consisting of homogeneous elements in R; then the relations

$$R(x_1, \cdots, x_i) = R \cap R_{\mathfrak{M}}(x_1, \cdots, x_i).$$

for $i = 1, \cdots, m$, show that a system of homogeneous elements $x_1, \cdots, x_m$ in R is a system of parameters for $\mathfrak{M}$ if and only if it is such for $R_{\mathfrak{M}}$, and that such a system of parameters for $\mathfrak{M}$ is distinct if and only if it is distinct as a system of parameters for $R_{\mathfrak{M}}$. It follows then from a well-known result for local ring (LR, 25.5) that if one system of homogeneous parameters for $\mathfrak{M}$ is distinct, then every system of homogeneous parameters for $\mathfrak{M}$ is distinct. We note incidentally that the definitions given here are compatible with the terminology in LR provided the elements $x_1, \cdots, x_m$ are homogeneous; without the homogeneity condition, our definitions would be somewhat stricter than those given in LR.

We shall be concerned in this paper only with the special case of an H-local ring R where $R_0 = K$ is a field; equivalently we can characterize this case by the condition that all homogeneous elements of positive degrees in R generate a maximal ideal $\mathfrak{M}$ in R, which must then be the only maximal homogeneous ideal in R. A simple but important consequence of this assumption is that, if m is the height of $\mathfrak{M}$, then there exist homogeneous elements $x_1, \cdots, x_m$ in $\mathfrak{M}$ such that the ideal $R(x_1, \cdots, x_r)$ has the height r; this implies in particular that $x_1, \cdots, x_m$ form a system of homogeneous parameters for $\mathfrak{M}$, since $\mathfrak{M}$ is the only homogeneous prime ideal in R having a height $\geqq m$. We shall prove this assertion by induction, assuming that the homogeneous elements $x_1, \cdots, x_r$ $(r < m)$ have already been chosen in $\mathfrak{M}$ so that $R(x_1, \cdots, x_r)$ has the height r. Let $\mathfrak{p}_1, \cdots, \mathfrak{p}_s$ be the minimal prime divisor of $R(x_1, \cdots, x_r)$ having the height r; we recall that they are all homogeneous ideals and hence are all contained in $\mathfrak{M}$. Consider any one ideal $\mathfrak{p}_i$; since $\mathfrak{p}_i$ does not contain $\bigcap_{j \neq i} \mathfrak{p}_j$ and since both ideals are homogeneous, there exists a homogeneous element z_i in $\bigcap_{j \neq i} \mathfrak{p}_j$ which is not contained in $\mathfrak{p}_i$. Let d_i be the degree of the homogeneous element z_i (observe that z_i as an element in $\mathfrak{M}$ has a positive degree), and set $c_i = d_1 \cdots d_s/d_i$; then the element $z_1^{c_1} + \cdots + z_s^{c_i}$ is a homogeneous element (of degree $d_1 \cdots d_s$) in $\mathfrak{M}$ and is not contained in any one of the ideal $\mathfrak{p}_1, \cdots, \mathfrak{p}_s$. If we take this element to be x_{r+1}, then clearly $R(x_1, \cdots, x_{r+1})$ must have the height $r+1$; this proves our assertion.

An H-local ring R is said to be H-Macaulay if its maximal homogeneous ideal $\mathfrak{M}$ has a distinct system of homogeneous parameters. The existence of a homogeneous system of parameters for $\mathfrak{M}$ in the special case $R_0 = K$ enables us to prove the following lemma; without the assumption $R_0 = K$, the "if" part still holds, but the "only if" part would be no longer generally true, as will be shown by a simple example below.

LEMMA 5. *Let* $R = \sum_{i=0}^{\infty} R_i$ *be a Notherian graded ring such that* R_0 *is* *a field; then the unmixedness theorem holds for homogeneous ideals in R if and only if R is H-Macaulay.*

Proof. If R is H-Macaulay, then $R_{\mathfrak{M}}$ is Macaulay and hence the unmixedness theorem hold in $R_{\mathfrak{M}}$; the relation $\mathfrak{a} = R \cap R_{\mathfrak{M}}\mathfrak{a}$ for any homogeneous ideal in R shows that then the unmixedness theorem holds for homogeneous ideals in R. Conversely, assume that the unmixedness theorem holds for homogeneous ideals in R, and let $x_1, \cdots, x_m$ be a system of homogeneous parameters for the maximal homogeneous ideal $\mathfrak{M}$ in R, with the property that $R(x_1, \cdots, x_i)$ has the height i for $i = 1, \cdots, m$; we have just noted above that such a system of homogeneous parameter exists. Since $R(x_1, \cdots, x_{i-1})$ has the height $i-1$ and does not have any embedded prime divisors, all its prime divisors must have the height $i-1$; it follows then that x_i is not contained in any of the prime divisors of $R(x_1, \cdots, x_{i-1})$ and hence is not a zero-divisor for the R-module $R/R(x_1, \cdots, x_{i-1})$. This shows that the system of homogeneous parameters $x_1, \cdots, x_m$ is distinct, so that R is H-Macaulay.

To show that an H-local ring does not necessarily have a system of homogeneous parameters for its maximal homogeneous ideal, even when the unmixedness theorem holds in it, consider the following example. Let $P[x] = P[x_1, \cdots, x_m]$ be the polynomial ring of m variables over a real discrete valuation ring P, graded as usual by assigning to each variable x_i the degree 1; it is clear that $P[x]$ is H-local, and that if t is a generator of the maximal ideal in P, then the maximal homogeneous ideal in $P[x]$ is generated by the $m+1$ elements $x_1, \cdots, x_m, t$. Furthermore, since the elements $x_1, \cdots, x_m, t$ clearly form a $P]x[$-sequence, they form a distinct system of homogeneous parameters for the maximal homogeneous ideal in $P[x]$, so that $P[x]$ is H-Macaulay. (In fact, it is well-known that $P[x]$ is locally Macaulay.) Now, let $\mathfrak{n}$ be the homogeneous ideal in $P[x]$ generated by the single element tx_m and consider the residue ring $R = P[x]/\mathfrak{n}$; it is clear that R as a graded ring is also H-local and that the maximal homogeneous ideal $\mathfrak{M}$ in R is generated by the images $\bar{x}_1, \cdots, \bar{t}_m, \bar{x}$ in R of $x_1, \cdots, x_m, t$ respectively. The relations $\bar{x}_m{}^2 = \bar{x}_m(\bar{x}_m + \bar{t})$ and $\bar{t}^2 = \bar{t}(\bar{x}_m + \bar{t})$ show that the ideal $R(\bar{x}_1, \cdots, \bar{x}_{m-1}, \bar{x}_m + \bar{t})$ is primary for $\mathfrak{M}$; on the other hand, since the ideal $R(\bar{x}_1, \cdots, \bar{x}_{m-1})$ has the height $m-1$ and since $\bar{x}_m + \bar{t}$ is not contained in either of the two prime divisors $R(\bar{x}_1, \cdots, \bar{x}_{m-1}, \bar{x}_m)$ and $R(\bar{x}_1, \cdots, \bar{x}_{m-1}, \bar{t})$ of $R(\bar{x}_1, \cdots, \bar{x}_{m-1})$, it follows that $\mathfrak{M}$ has the height m and hence the elements $\bar{x}_1, \cdots, \bar{x}_{m-1}. \bar{x}_m + \bar{t}$ form a system of parameters for M. In fact, our argument shows that $\bar{x}_m + \bar{t}$ is not a zero-divisor for the

R-module $R/R(\bar{x}_1,\cdots,\bar{x}_{m-1})$, and since $\bar{x}_i$ is evidently not a zero-divisor for the R-module $R/R(\bar{x}_1,\cdots,\bar{x}_{i-1})$ for $i=1,\cdots,m-1$, it follows that $\bar{x}_1,\cdots,\bar{x}_{m-1},\bar{x}_m+\bar{t}$ form a distinct system of parameters for $\mathfrak{M}$. This implies that $R_{\mathfrak{M}}$ is Macaulay and hence the unmixedness theorem holds in $R_{\mathfrak{M}}$; and the basic relation $\mathfrak{a}=R\cap R_{\mathfrak{M}}\mathfrak{a}$ for any homogeneous ideal in R shows then that the unmixedness theorem holds for homogeneous ideals in R. However, since the element $\bar{x}_m+\bar{t}$ is not homogeneous, the system $\bar{x}_1,\cdots,\bar{x}_{m-1}$, $\bar{x}_m+\bar{t}$ is not a system of homogeneous parameters for $\mathfrak{M}$; in fact, we maintain that such a system of parameters cannot exist. For any system of homogeneous parameters for $\mathfrak{M}$ cannot consist of positive degrees only, for then all elements in the system would be contained in $R(\bar{x}_1,\cdots,\bar{x}_m)$, which has the height $m-1$. On the other hand, if such a system contains a homogeneous element of degree 0 in $\mathfrak{M}$, the image system in the residue ring $R/R(\bar{t})$ contains less than m non-zero elements and yet must generate an ideal which is primary for the ideal $\mathfrak{M}/R(t)$; but this is impossible, for $R/R(t)$ is isomorphic to the polynomial ring of m variables over the field $P/P(t)$ and in this isomorphism the ideal $\mathfrak{M}/R(t)$ corresponds to the ideal generated by the m independent variables.

Let $R=\sum_{i=0}^{\infty}R_i$ and $S=\sum_{i=0}^{\infty}S_i$ be two Noetherian graded rings which both contain the same field K and satisfy the condition $R_0=S_0=K$; in other words, R and S are Noetherin graded K-algebras which both contain K as the subring of all homogeneous elements of degree 0. Let $\mathfrak{M}$ and $\mathfrak{N}$ be the maximal homogeneous ideals in R and S respectively, which we recall are the ideals generated by all homogeneous elements of positive degrees in R and S respectively. Consider the tensor product $T=R\otimes S$ of R and S over K; as usual, we shall consider both R and S embedded canonically as subalgebras in T, so that the tensor product between elements in R and S can be expressed as ordinary product in T. The gradings of R and S induce in a natural way a bi-grading of $T=\sum_{i,j=0}^{\infty}T_{i,j}$, where $T_{i,j}$ is the K-module generated by all products of homogeneous elements of degree i in R with homogeneous elements of degree j in S. An element in $T_{i,j}$ will be said to be homogeneous of bi-degree (i,j), and the bi-degree (i,j) is said to be positive if either i or j is positive; we note that $T_{0,0}=K$ and that $\mathfrak{D}=(\mathfrak{M},\mathfrak{N})$ is the maximal prime ideal in T generated by all homogeneous elements of positive bi-degrees. Furthermore, we can make T into a graded ring by assigning to each element in $T_{i,j}$ the degree $i+j$, so that we can set $T=\sum_{i=0}^{\infty}T_i$ with $T_i=\sum_{j+k=i}T_{j,k}$; and we note again that $T_0 = K$ and that $\mathfrak{O}$ is the ideal generated in T by all homogeneous elements of positive degrees.

LEMMA 6. *If $a_1, \cdots, a_r$ is an R-sequence in R and $b_1, \cdots, b_s$ is an S-sequence in S, then $a_1, \cdots, a_r, b_1, \cdots, b_s$ is a T-sequence in T.*

Proof. We use induction on the total number $r + s$, the assertion being trivially true in case $r = s = 0$. Since by induction hypothesis the system $a_1, \cdots, a_r, b_1, \cdots, b_{s-1}$ is already a T-sequence in T, we have only to show that b_s is not a zero-divisor for the T-module $T/T(a_1, \cdots, a_r, b_1, \cdots, b_{s-1})$. Now, there is a ring isomorphism between $T/T(a_1, \cdots, a_r, b_1, \cdots, b_{s-1})$ and the tensor product $(R/R(a_1, \cdots, a_r)) \otimes (S/S(b_1, \cdots, b_{s-1}))$, and the image of b_s in $S/S(b_1, \cdots, b_{s-1})$ under the canonical mapping is by assumption not a zero-divisor; our assertion follows then from the fact that in the tensor product of two algebras over a field any regular element in one of the algebras cannot become a zero-divisor.

An almost immediate consequence of Lemma 6 is that if both R and S are H-Macaulay, then the tensor product T is also H-Macaulay. In fact, if the system $x_1, \cdots, x_m$ is a distinct system of homogeneous parameters for $\mathfrak{M}$ and the system $y_1, \cdots, y_n$ is a distinct system of homogeneous parameters for $\mathfrak{N}$, then the combined system $x_1, \cdots, x_m, y_1, \cdots, y_n$ is by Lemma 6 a T-sequence in T; and since this combined system evidently generate in T an ideal which is primary for the ideal $\mathfrak{O} = T(\mathfrak{M}, \mathfrak{N})$, it is a distinct system of homogeneous parameters for $\mathfrak{O}$. Geometrically, this implies that if the unmixedness theorem holds for two varieties U and V, then it holds also for their join $J(U, V)$.

As mentioned before, we are interested here not so much in the tensor product as in what we shall call the Segre product of R and S over K. For any pair of positive integers d and e, we define the Segre product of order (d, e) of R and S over K to be the Noetherian graded subalgebra $A = \sum_{i=0}^{\infty} A_i = \sum_{i=0}^{\infty} T_{id,ie}$ in the tensor product T; for brevity we shall simply speak of the Segre product of R and S in case the base field K as well as the order (d, e) are clear from the context. We observe that the grading of A is induced not only by the grading of T as a graded ring $T = \sum_{i=0}^{\infty} T_i$ but also by the bi-grading of T as a bi-graded ring $T = \sum_{i,j=0}^{\infty} T_{i,j}$, only in the latter case the induced bi-grading of A becomes a simple grading on account of the absence of any elements (except 0) of bi-degrees different from the form (id, ie). It is clear that $A_0 = K$, and that the ideal $\mathfrak{L} = A \cap \mathfrak{O}$ is the ideal in A generated by all homogeneous elements of positive degrees and hence is the maximal homogeneous ideal in A.

We wish to construct a system of parameters for the ideal $\mathfrak{L}$ consisting of homogeneous elements of the same degree in A. As we have shown before, there exists a system of homogeneous parameters for the ideal $\mathfrak{M}$ in R; moreover, by replacing each element in such a system by a suitable power of itself, we can obtain a system of parameters for $\mathfrak{M}$ in which every element is homogeneous of the degree fd, where f can be any multiple of a suitable positive integer. Similarly, we can obtain a system of parameters for $\mathfrak{N}$ consisting of homogeneous elements of degree fe in S, where f again can be any multiple of a suitable positive integer. We now take for f an arbitrary but fixed integer which satisfies both these conditions; let $x_1, \cdots, x_m$ be a system of parameters for $\mathfrak{M}$ consisting of homogeneous elements of degree fd in R, and let $y_1, \cdots, y_n$ be a system of parameters for $\mathfrak{N}$ consisting of homogeneous elements of degree fe in S. Next, let $u_1, \cdots, u_{m+n-1}$ be linear combinations in the $x_1, \cdots, x_m$ with coefficients in K such that every m elements among them are linearly independent (as linear forms over K); it is easily seen that then every m elements among the $u_1, \cdots, u_{m+n-1}$ form also a system of parameters for $\mathfrak{M}$. Similarly, let $v_1, \cdots, v_{m+n-1}$ be linear combinations in the $y_1, \cdots, y_n$ with coefficients in K such that every n elements among them are linearly independent (as linear forms over K); then every n elements among the $v_1, \cdots, v_{m+n-1}$ form also a system of parameters for $\mathfrak{N}$. In order to insure the existence of these linear combinations, we shall assume that the field K contains sufficiently large number of elements (for fixed integers m and n), e. g., K is infinite; however, this restriction is not serious, for not only can it often be met in applications by a change of the ground field, but also it can be removed in the general case by the usual trick of adjunction of a variable to the field K. Now, we set $w_i = u_i v_i$, $i = 1, \cdots, m + n - 1$; it is clear that $w_1, \cdots, w_{m+n-1}$ are all bi-homogeneous elements of bi-degree (fd, fe) in T and hence are homogeneous elements of degree f in A. We shall show presently that $w_1, \cdots, w_{m+n-1}$ form a system of parameters for $\mathfrak{L}$. Before doing this, we shall prove a few lemmas concerning the relations between the ideals in T and the ideals in A.

LEMMA 7. *For any ideal* $\mathfrak{a}$ *in* A, *we have the relation* $A \cap T\mathfrak{a} = \mathfrak{a}$.

Proof. Let B be the K-submodule in T generated by all homogeneous elements of bi-degrees not of the type (ld, le) for any l; then we have the direct sum $T = A + B$ as K-modules, and it is easily seen that B is also an A-module. Now, let $a_1, \cdots, a_r$ be a base of the ideal $\mathfrak{a}$, and let $z = \sum_{i=1}^{r} c_i a_i$ be an element in $A \cap T\mathfrak{a}$, where the c_i are elements in T. If we set $c_i = c_i' + c_i''$, with c_i' in A and c_i'' in B, then we have the relation $z = \sum_{i=1}^{r} c_i' a_i + \sum_{i=1}^{r} c_i'' a_i$; since $\sum_{i=1}^{r} c_i'' a_i$ is an element in both A and B, it must be equal to 0. Therefore $z = \sum_{i=1}^{r} c_i' a_i$, which shows that z is in $\mathfrak{a}$.

LEMMA 8. *If* $\mathfrak{P}$ *is a bi-homogeneous prime ideal in* T *which does not contain either* $T\mathfrak{M}$ *or* $T\mathfrak{N}$, *then* $\mathfrak{P}$ *is the radical of* $T(A \cap \mathfrak{P})$.

Proof. Since $\mathfrak{P}$ does not contain either $T\mathfrak{M}$ or $T\mathfrak{N}$, it cannot contain all the x_i or all the y_i; we can therefore assume that, say, x_j and y_k are not in P. Let E be the multiplicatively closed set generated by x_j and y_k, i.e., E is the set of all monomials in x_j, y_k; since E does not intersect $\mathfrak{P}$, we have the usual relations of a quotient ring, $T \cap T_E\mathfrak{P} = \mathfrak{P}$ and $T \cap (T_E(A \cap \mathfrak{P})) = T(A \cap \mathfrak{P})$. Let $p_1, \cdots, p_s$ be a base for $\mathfrak{P}$ consisting of bi-homogeneous elements, and let $(g_1, h_1), \cdots, (g_s, h_s)$ be their bi-degrees respectively; then, using the notations introduced above, the ideal $A \cap \mathfrak{P}$ contains the elements $p_i^{def} x_j^{dh_i} y_k^{eg_i}$, $i = 1, \cdots, s$, for they have the bi-degrees $(df(eg_i + dh_i), ef(dh_i + eg_i))$, $i = 1, \cdots, s$, respectively. Then it is easily seen that $T_E(A \cap \mathfrak{P})$ must contain the elements p_i^{def}, $i = 1, \cdots, s$, and hence $T(A \cap \mathfrak{P}) = T \cap (T_E(A \cap \mathfrak{P}))$ must also contain these s elements. This shows that $\mathfrak{P}$ is contained in the radical of $T(A \cap \mathfrak{P})$ and hence must coincide with it.

LEMMA 9. *If* $\mathfrak{P}$ *is a bi-homogeneous prime ideal in* T *which does not contain either* $T\mathfrak{M}$ *or* $T\mathfrak{N}$, *then* $\mathfrak{P}$ *and* $A \cap \mathfrak{P}$ *have the same height.*

Proof. Any chain of distinct prime ideals $\mathfrak{P} \supset \mathfrak{P}_1 \supset \cdots \supset \mathfrak{P}_t$ will contract in A to a chain of prime ideals $A \cap \mathfrak{P} \supset A \supset \mathfrak{P}_1 \supset \cdots \supset A \cap \mathfrak{P}_t$, and Lemma 8 shows that the ideals in the latter chain must be distinct; hence the height of $\mathfrak{P}$ is not greater than that of $A \cap \mathfrak{P}$. To prove the converse, consider first a prime ideal $\mathfrak{p}$ in A; it follows easily from Lemma 7 that there must be at least one minimal prime divisor $\mathfrak{P}'$ of $T\mathfrak{p}$ which contracts in A to $\mathfrak{p}$. By Lemma 8, $\mathfrak{P}'$ coincides with the radical of $T\mathfrak{p}$ and hence must be the only minimal prime divisor of $T\mathfrak{p}$. Now, if $A \cap \mathfrak{P} \supset \mathfrak{p}_1 \supset \cdots \supset \mathfrak{p}_t$ is a chain of distinct prime ideals in A, then we can obtain a chain of prime ideals $\mathfrak{P} \supset \mathfrak{P}_1 \supset \cdots \supset \mathfrak{P}_t$ in T such that each $\mathfrak{P}_i$ is a minimal prime divisor of $T\mathfrak{p}_i$; it follows from what we have just shown that each $\mathfrak{P}_i$ is the unique minimal prime divisor of $T\mathfrak{p}_i$ which contracts in A to $\mathfrak{p}_i$, and hence the ideals in the latter chain must be distinct. This shows that the height of $A \cap \mathfrak{P}$ is not greater than that of $\mathfrak{P}$.

LEMMA 10. *The elements* $w_1, \cdots, w_{m+n-1}$ *form a system of homogeneous parameters for maximal homogeneous ideal* $\mathfrak{L}$ *in* A.

Proof. We can choose the system $x_1, \cdots, x_m$ so that $R(x_1, \cdots, x_r)$ has the height r for every $r \leqq m$, and choose the system $y_1, \cdots, y_n$ so that $S(y_1, \cdots, y_s)$ has the height s for every $s \leqq n$; then it can be shown that $T(x_1, \cdots, x_r, y_1, \cdots, y_s)$ has the height $r+s$ (the proof is simple, but we shall not go into details here). In particular, the ideal $T(x_1, \cdots, x_{m-1}, y_1, \cdots, y_{n-1})$ has the height $m+n-2$, and hence (LR, 9.3) every one of its minimal prime divisors must have the height $m+n-2$. If $\mathfrak{P}$ is any one such prime divisor, then $\mathfrak{P}$ cannot contain the element $x_m y_n$; for, otherwise $\mathfrak{P}$ must contain either x_m or y_n, which would imply that either $T(x_1, \cdots, x_m, y_1, \cdots, y_{n-1})$ or $T(x_1, \cdots, x_{m-1}, y_1, \cdots, y_n)$ must have the height $m+n-2$. Since we can choose the u's and the v's so that $w_1 = u_1 v_1$ (or any other w_i) is equal to $x_m y_n \bmod T(x_1, \cdots, x_{m-1}, y_1, \cdots, y_{n-1})$, it follows that $\mathfrak{P}$ cannot contain w_1 and hence the ideal $A \cap \mathfrak{P}$ is a proper subset in $\mathfrak{L}$. Now, since $\mathfrak{P}$ does not contain either $T\mathfrak{M}$ or $T\mathfrak{N}$, Lemma 9 shows that $A \cap \mathfrak{P}$ must have the same height $m+n-2$ as $\mathfrak{P}$ itself; this shows that the height of $\mathfrak{L}$ must be at least $m+n-1$. On the other hand, it is clear that the system $u_1, \cdots, u_{m+n-1}$ and $v_1, \cdots, v_{m+n-1}$ introduced in this section satisfy the condition (B) in Section 2 (with $t = m+n-1$); hence we can apply Lemma 4, by the Remark after that lemma. The application of Formula (6) in Lemma 4 to our present situation shows that

$$T(w_1, \cdots, w_{m+n-1}) = T(x) \cap T(y) \cap T(w, (x, y)^{m+n-1});$$

hence, by Lemma 7, we have the decomposition

$$\begin{aligned} A(w_1, \cdots, w_{m+n-1}) &= A \cap T(w_1, \cdots, w_{m+n-1}) \\ &= (A \cap T(x)) \cap (A \cap T(y)) \cap (A \cap T(w, (x, y)^{m+n-1})). \end{aligned}$$

Since it can be easily seen that a sufficiently high power of every element in $\mathfrak{L}$ is contained in every one of the three ideals appearing in the expression on the right hand side of this equation, it follows that $A(w_1, \cdots, w_{m+n-1})$ is primary for $\mathfrak{L}$. This shows that the height of $\mathfrak{L}$ is at most $m+n-1$ and hence must be equal to $m+n-1$; therefore the elements $w_1, \cdots, w_{m+n-1}$ form a system of homogeneous parameters for $\mathfrak{L}$.

4. Our aim is to investigate the validity of the unmixedness theorem for homogeneous ideals in the Segre product A of R and S, assuming that the unmixedness theorem holds for homogeneous ideals in both R and S. It turns out that a certain additional condition on both R and S has to be introduced, which proves to be both necessary and sufficient (except in some very special cases). In order to be able to formulate our result properly, we shall first make a few remarks about this condition.

As before, let $R = \sum_{i=0}^{\infty} R_i$ be a Noetherian graded ring with $R_0 = K$, and let $\mathfrak{M}$ be the ideal generated in R by all elements of positive degrees. For simplicity, we shall call a system of parameters for $\mathfrak{M}$ *homogeneous* if it consists of homogeneous elements of the same degree in R, and this common degree is then called the *degree* of the system. Let $x_1, \cdots, x_m$ be a homogeneous system of parameters of degree fd for M; we shall say that this system has the d-exponent ϵ (or simply the exponent ϵ in case $d = 1$) if ϵ is the smallest integer with the property that every homogeneous element in R of degree ld with $l \geqq \epsilon f$ is contained in $R(x_1, \cdots, x_m)$. In case R is a homogeneous ring (i.e., $\mathfrak{M}$ is generated by homogeneous elements of degree 1) and $fd = 1$, then the exponent of the homogeneous system of parameters $x_1, \cdots, x_m$ is simply the exponent of the primary ideal $R(x_1, \cdots, x_m)$. We shall show that, under the assumption that R is H-Macaulay, if any one homogeneous system of parameters of degree fd for $\mathfrak{M}$ has the d-exponent ϵ, then every homogeneous system of parameters of degree fd will have the same d-exponent ϵ. It is sufficient to show that if the system $x_1, \cdots, x_m$ has d-exponent $\leqq \epsilon$, then the same is true of the homogeneous system of parameters $x_1, \cdots, x_{m-1}, z$ obtained from it by replacing x_m by another homogeneous element of degree fd in R. Let a be a homogeneous of degree ld in R with $l \geqq \epsilon f$; since the ideal $R(x_1, \cdots, x_{m-1}, z)$ is primary for $\mathfrak{M}$, there is a positive integer β such that $a x_m^{\beta}$ is contained in $R(x_1, \cdots, x_{m-1}, z)$. We have then a relation of the form $a x_m^{\beta} = a_1 x_1 + \cdots + a_{m-1} x_{m-1} + a_m z$, where $a_1, \cdots, a_m$ are homogeneous elements of degree $(l + \beta f - f) d$ in R. Since $(l + \beta f - f) \geqq \epsilon f$, we have the relation $a_m = b_1 x_1 + \cdots + b_m x_m$, by our assumption about the d-exponent of the system $x_1, \cdots, x_m$. Combining these two relations, we obtain the relation $(a_1 + b_1 z) x_1 + \cdots + (a_{m-1} + b_{m-1} z) x_{m-1} + (b_m z - a x_m^{\beta-1}) x_m = 0$, from which we conclude that $b_m z - a z_m^{\beta-1}$ is contained in $R(x_1, \cdots, x_{m-1})$ and hence $a x_m^{\beta-1}$ is contained in $R(x_1, \cdots, x_{m-1}, z)$. We can then repeat the same argument for $a x_m^{\beta-1}$, and conclude after β steps that a itself is contained in $R(x_1, \cdots, x_{m-1}, z)$.

It is easily seen that the d-exponent does not in general remain unchanged if we replace a homogeneous system of parameters for $\mathfrak{M}$ of degree fd by another one of degree $f'd$ with a different f'. On the other hand, we shall show that, always under the assumption that R is H-Macaulay, if any one (and hence every) homogeneous system of parameters of degree fd for $\mathfrak{M}$ has a d-exponent $\leqq m$, then the same is true for every homogeneous system of parameters of degree $f'd$ for any f'. It is clearly sufficient to show this for the case where either f is a multiple of f' or f' is a multiple of f; futhermore, in view of what we have just shown above, it is sufficient to show this for any two particular homogeneous systems of parameters of degrees fd and $f'd$. It is therefore sufficient to show that the system $x_1,\cdots,x_m$ has a d-exponent $\leqq m$ if and only if the system $x_1^s,\cdots,x_m^s$ (for any positive integer s) has a d-exponent $\leqq m$.

Assume first that the system $x_1,\cdots,x_m$ has a d-exponent $\leqq m$, and let a be a homogeneous element of degree ld with $l\geqq msf$. Since evidently $l\geqq mf$ we have the relation $a=a_1x_1+\cdots+a_mx_m$, where $a_1,\cdots,a_m$ are homogeneous elements of degree $(l-f)d$; since $l-f\geqq(s-1)mf\geqq mf$, the elements $a_1,\cdots,a_m$ are themselves contained in $R(x_1,\cdots,x_m)$. Thus proceeding in this way, we conclude after $\delta=m(s-1)+1$ steps that $a=\sum a_{i_1\cdots i_\delta}x_{i_1}\cdots x_{i_\delta}$. Since every term in this expression must contain the s-th power of at least one of the x_i, it follows that a is contained in $R(x_1^s,\cdots,x_m^s)$. Conversely, assume that the system $x_1^s,\cdots,x_m^s$ has a d-exponent $\leqq m$, and let a be a homogeneous element of degree ld with $l\geqq mf$. Consider the product $a(x_1\cdots x_m)^{s-1}$, which is a homogeneous element of degree $(l+(s-1)mf)d$; since $l+(s-1)mf\geqq msf$, this element $a(x_1\cdots x_m)^{s-1}$ must be contained in $R(x_1^s,\cdots,x_m^s)$, so that we have the $a(x_1\cdots x_m)^{s-1}=\sum_{i=1}^{m}a_ix_i^s$. If a were not contained in $R(x_1,\cdots,x_m)$, then this relation could be written as a non-trivial relation in the $x_1,\cdots,x_m$ with coefficients which are not in $R(x_1,\cdots,x_m)$; in fact this relation would not be trivial (i.e., both sides being identical as forms in the $x_1,\cdots,x_m$) because the left side could contain no s-th power of any one of the $x_1,\cdots,x_m$ while each term in the right side would contain the s-th power of at least one of them. Since such a relation cannot exist (LR, 25.2), we conclude that a must be contained in $R(x_1,\cdots,x_m)$.

Thus we have shown that the property of having a d-exponent $\leqq m$ is independent of the choice of the homogeneous system of parameters for $\mathfrak{M}$, (provided of course it has a degree which is a multiple of d) and hence can be considered as a property of the H-Macaulay graded ring R itself. We shall express this property by saying that the H-Macaulay ring R is *d-proper*, or simply *proper* in case $d=1$. It is easily seen from our above argument that if R is d-proper, then it is also d'-proper for any d' which is a multiple of d; in particular, if R is proper, then it is d-proper for every d.

We are now in a position to prove our main theorem.

THEOREM. *Let R and S be Noetherian graded K-algebras with $R_0 = S_0 = K$, and let m and n be the heights of the prime ideals generated by all the homogeneous elements of positive degrees in R and S respectively; we assume that both R and S are H-Macaulay. Then the Segre Product A of order (d, e) of R and S over K is H-Macaulay if R is d-proper* (*unless* $n = 1$) *and S is e-proper* (*unless* $m = 1$); *and this condition is also necessary in case R and S contain non-zero homogeneous elements of all degrees which are multiples of d and e respectively. Furthermore, when R and S are respectively d-proper and e-proper, then the H-Macaulay ring A is proper.*

Proof. Using the results and notations developed in Section 3, our theorem, except for the last assertion, is equivalent to the statement that the system of parameters $w_1, \cdots, w_{m+n-1}$ for $\mathfrak{L}$ is distinct if and only if the stated condition is satisfied. Now, as we have observed before, the system of parameters $w_1, \cdots, w_{m+n-1}$ is distinct if and only if the ideal $A(w_1, \cdots, w_t)$ is unmixed for every $t = 1, \cdots, m+n-2$. By Lemma 7, we have the relation $A(w_1, \cdots, w_t) = A \cap T(w_1, \cdots, w_t)$; therefore, if $T(w_1, \cdots, w_t) = \bigcap_i \mathfrak{Q}_i$ is a (not necessarily shortest) primary decomposition of $T(w_1, \cdots, w_t)$, i.e., a representation of $T(w_1, \cdots, w_t)$ as an intersection of primary ideals, then $A(w_1, \cdots, w_t) = \bigcap_i (A \cap \mathfrak{Q}_i)$ is a primary decomposition of $A(w_1, \cdots, w_t)$. In order that $A(w_1, \cdots, w_t)$ be unmixed, it is therefore necessary and sufficient that every ideal $A \cap \mathfrak{Q}_i$ which is not redundant in the primary decomposition has the height t.

We now apply the results of Section 2 to the ring T, and observe that the systems $u_1, \cdots, u_t$ and $v_1, \cdots, v_t$, as constructed in Section 3, satisfy the conditions stated in Section 2. We note here a small discrepancy in convention between Section 2 and our present situation, namely in Section 2 the integer t is always not less than m or n, while now t may be any integer from 1 to $m+n-2$. However, this discrepancy in convention can be easily taken care of; for, in case t is less than m or n, we can replace m or n respectively by t, and one verifies easily that this change of m or n will not invalidate the inequality $t \leqq m+n-2$ except when $n = 1$ or $m = 1$ respectively (after the change in case both m and n are changed). Therefore we can apply formula (1) of Lemma 3 to our present situation for all values of t from 1 to $m+n-2$:

$$(7)\quad \begin{aligned} &T(w_1,\cdots,w_t) = \\ &\quad \bigcap_{(i)} T(u_{i_1},\cdots,u_{i_\alpha}, v_{i_{\alpha+1}},\cdots,v_{i_t}) \cap T(u_1,\cdots,u_t) \cap T(v_1,\cdots,v_t), \end{aligned}$$

where the intersection extends over all partitions (i) of the integers from 1 to t into two sets $(i_1,\cdots,i_\alpha)$ and $(i_{\alpha+1},\cdots,i_t)$, for all values of α from $\sup(t-n+1,0)$ to $\inf(m-1,t)$. We observe that in case $t<m$ the term $T(u_1,\cdots,u_t)$ no longer coincides with $T(u_1,\cdots,u_m)=T(x_1,\cdots,x_m)$ and is in fact redundant in (7); a similar statement holds of course for the term $T(v_1,\cdots,v_t)$ in case $t<n$.

Now, since each ideal $T(u_{i_1},\cdots,u_{i_\alpha}, v_{i_{\alpha+1}},\cdots,v_{i_t})$ is generated by t elements in a distinct system of parameters for $\mathfrak{O}$, it has the height t. It follows then from the validity of the unmixedness theorem for homogeneous ideals in T that this ideal has a (shortest) primary decomposition

$$T(u_{i_1},\cdots,u_{i_\alpha}, v_{i_{\alpha+1}},\cdots,v_{i_t}) = \bigcap_{j=1}^{\alpha(i)} \mathfrak{Q}_{(i),j},$$

in which every ideal $\mathfrak{Q}_{(i),j}$ has the height t. Furthermore, since this is a bi-homogeneous ideal in T, every ideal $\mathfrak{Q}_{(i),j}$ is also bi-homogeneous. If we assume for convenience $m\leqq n$ and set $T(x)=T(x_1,\cdots,x_m)$ and $T(y)=T(y_1,\cdots,y_n)$, then

$$\begin{aligned} T(w_1,\cdots,w_t) &= \bigcap_{(i)} \bigcap_{j=1}^{\alpha(i)} \mathfrak{Q}_{(i),j} && (t<m), \\ &= \bigcap_{(i)} \bigcap^{\alpha(i)} \mathfrak{Q}_{(i),j} \cap T(x) && (m\leqq t<n), \\ &= \bigcap_{(i)} \bigcap_{j=1}^{\alpha(i)} \mathfrak{Q}_{(i),j} \cap T(x) \cap T(y) && (n\leqq t). \end{aligned}$$

If we set $\mathfrak{q}_{(i),j}=A\cap\mathfrak{Q}_{(i),j}$, then it follows from the relation $A(w_1,\cdots,w_t)=A\cap T(w_1,\cdots,w_t)$ that

$$(8)\quad \begin{aligned} A(w_1,\cdots,w_t) &= \bigcap_{(i)} \bigcap_{j=1}^{\alpha(i)} \mathfrak{q}_{(i),j} && (t<m), \\ &= \bigcap_{(i)} \bigcap_{j=1}^{\alpha(i)} \mathfrak{q}_{(i),j} \cap (A\cap T(x)) && (m\leqq t<n), \\ &= \bigcap_{(i)} \bigcap_{j=1}^{\alpha(i)} \mathfrak{q}_{(i),j} \cap (A\cap T(x)) \cap (A\cap T(y)) && (n\leqq t). \end{aligned}$$

Since all the ideals $\mathfrak{Q}_{(i),j}$ are bi-homogeneous and none of them contains either $T\mathfrak{M}$ or $T\mathfrak{N}$, it follows from Lemma 9 that the contraction $\mathfrak{q}_{(i),j}$ in A

of every $\mathfrak{Q}_{(i),j}$ has the height t. On the other hand, it is clear that both $A \cap T(x)$ and $A \cap T(y)$ are primary ideals for $\mathfrak{L}$, for a sufficiently high power of every element in $\mathfrak{L}$ is easily seen to be contained in both $A \cap T(x)$ and $A \cap T(y)$. It follows then that $A(w_1, \cdots, w_t)$ is unmixed if and only if the ideal $A \cap T(x)$ and $A \cap T(y)$, whenever either or both occur in the above primary decomposition of $A(w_1, \cdots, w_t)$, are in effect redundant. We have therefore to prove that this is so if and only if the stated condition is satisfied.

Consider the ideal $A \cap T(x)$. If $n = 1$, then $t < m$ and hence $A \cap T(x)$ does not appear in the right side of (8), so that there is nothing to prove. Assume that $n > 1$, and take $t \geqq m$ (which is possible in case $n > 1$); let z be any element in A which is contained in every one of the ideals $\mathfrak{q}_{(i),j}$. Since z is evidently contained in every one of the ideals $T(u_{i_1}, \cdots, u_{i_a}, v_{i_{a+1}}, \cdots, v_{i_t})$ appearing in the expression on the right side of (7), it follows from Lemma 1 that z is mod $T(x)$ equal to a sum of elements of the form $cv_{i_1} \cdots v_{i_m}$. Furthermore, since z is contained in A, each such element $cv_{i_1} \cdots v_{i_m}$ must itself be a sum of terms of the form $abv_{i_1} \cdots v_{i_m}$, where a is a homogeneous element of degree ld in R with $l \geqq mf$, and b is a homogeneous element of degree $(l - mf)e$ in S. Conversely, if a is any homogeneous element of degree ld in R with $l \geqq mf$, and if we take for b any homogeneous element of degree $(l - mf)e$ in S, then the element $abv_{i_1} \cdots v_{i_m}$ is contained in every ideal $T(u_{i_1}, \cdots, u_{i_a}, v_{i_{a+1}}, \cdots, v_{i_t})$ appearing in the the right side of (7); and since this element has the bi-degree (ld, le), it is contained in A and hence is contained in every one of the ideals $\mathfrak{q}_{(i),j}$. Therefore, the ideal $A \cap T(x)$ is redundant in (8) if and only if for every homogeneous element a of degree ld in R with $l \geqq mf$, the element $abv_{i_1} \cdots v_{i_m}$ is contained in $T(x)$, for all choices of b and $v_{i_1}, \cdots, v_{i_m}$. We recall now that, for any ideal $\mathfrak{a}$ in R, the product of an element a in R with a non-zero element in S is contained in $T\mathfrak{a}$ if and only if a itself is contained in $\mathfrak{a}$; in fact, in view of the canonical isomorphism $T/T\mathfrak{a} = (R/\mathfrak{a}) \otimes S$, this follows from the well-known fact that the product in T of a non-zero element in R with a non-zero-element in S cannot be zero. If we apply this to the ideal $R(x)$ in R and observe that $bv_{i_1} \cdots v_{i_m}$ is a non-zero element in S, we conclude that $A \cap T(x)$ is redundant in (8) if and (in case S contains non-zero homogeneous elements of all degrees which are multiples of e) only if every homogeneous element of degre ld in R with $l \geqq mf$ is containd in $R(x)$, i. e., R is d-proper. Similarly, we can show that, unless $m = 1$ (in which case there is no condition), the ideal $A \cap T(y)$ is redundant in (8) if and (in case R contains non-zero homogeneous elements of all degrees which are multiples of d) only if S is e-proper. This jroves our theorem except for the last assertion.

To prove the last assertion, consider a homogeneous element z of degree l in A with $l \geqq (m+n-1)f$; since z as an element in T has the bi-degree (ld, le) and hence is a sum of elements of the type ab, where a and b are homogeneous elements of degrees ld and le in R and S respectively, it is sufficient to show every such element ab is contained in $T(w_1, \cdots, w_{m+n-1})$. Since R is d-proper and we have the inequality $l \geqq mf$, the element a is contained in $R(x)$ and hence can be exprssed in the form $a = \sum_i a_i x_i$ where the a_i are homogenous elements of degree $(l-f)d$ in R. If $n=1$, then we stop here; if $n>1$, then $l-f \geqq mf$ and hence each a_i can be expressed in the form $a_i = \sum_j a_{ij} x_j$, so that we can set $a = \sum_{i,j} a_{ij} x_j x_i$, where the a_{ij} are homogeneous elements of degree $(l-2f)d$ in R. Thus proceeding in this manner we obtain after n steps an expression for a of the form

$$a = \sum_{i_1, \cdots, i_n} a_{i_1 \cdots i_n} x_{i_1} \cdots x_{i_n}.$$

Similarly, we can obtain for the element b an expression of the form

$$b = \sum_{j_1, \cdots, j_m} b_{j_1 \cdots j_m} y_{j_1} \cdots y_{j_m}.$$

Then the element ab can be expressed as the sum of elements each of which is a monomial of degre $m+n$ in the x's and y's; hence, ab is contained in the ideal $T(w, (u, v)^{m+n})$, in the notation of Lemma 4. Since ab is evidently contained in both $T(x) = T(u)$ and $T(y) = T(v)$, it follows from Lemma 4 that ab is contained in $T(w_1, \cdots, w_{m+n-1})$.

We conclude this paper by giving a few samples of the application of our result. Let $R = K[x_1, \cdots, x_m]$ be a polynomial ring generated by m variables $x_1, \cdots, x_m$ over a field K. The usual way of making R into a graded ring is to assign each variable x_i the degree 1; we shall generalize the the situation here by assigning each variable x_i an arbitrary positive integer d_i as its degree, so that a "monomial" of the form $x_1^{r_1} \cdots x_m^{r_m}$ is a homogeneous element of degree $r_1 d_1 + \cdots + r_m d_m$ in R. It is clear that $R_0 = K$ in the so graded ring R and that m is the height of the maximal prime ideal $\mathfrak{M}$ generated by all homogeneous elements of positive degrees in R. Furthermore, it is well-known (LR, 25.10) that R is locally Macaulay and hence is of course also H-Macaulay, Finally, R is d-proper for any positive integer d. To see this, let $f = d_1 \cdots d_m$, and $f_i = \prod_{j \neq i} d_j$, and set $u_i = x_i^{df_i}$, so that $u_1, \cdots, u_m$ form a homogeneous system of parameters of degree fd for $\mathfrak{M}$; it is sufficient to show that an element of the form $z = x_1^{r_1} \cdots x_m^{r_m}$ of degree $ld = r_1d_1 + \cdots + r_md_m$ with $l \geqq mf$ is contained in the ideal $R(u_1, \cdots, u_m)$. Now, it is clear that the inequality $r_1d_1 + \cdots + r_md_m \geqq mfd$ implies the inequality $r_id_i \geqq fd$ for at least one i, and hence the inequality $r_i \geqq f_id$ for at least one i; and this implies that $x_i^{r_i}$ is contained in $R(u_i)$ and hence z is contained in $R(u_i)$. Thus we have shown that the graded ring R satisfies the hypothesis of our main theorem. If we take for S a polynomial ring $S = K[y_1, \cdots, y_n]$ generated by n variables $y_1, \cdots, y_n$ over K, graded by assigning each variable y_j an arbitrary positive integer e_j as its degree, and let e be any positive integer, then our main theorem shows that the Segre product A of order (d, e) of R and S is a proper H-Macaulay ring.

Finally, we give a couple of simple examples where the condition in our main theorem is not satisfied. Let $R = K[x_1, x_2, x_3]$ be the homogeneous coordinate ring of the plane curve defined by the equatoon $x_3^3 - x_1^2x_2 = 0$. It is clear that x_1, x_2 form a distinct system of homogeneous parameters for the maximal homogeneous ideal $\mathfrak{M}$ in R. Since any relation between x_1, x_2, x_3 must be a multiple of the basic relation $x_3^3 - x_1^2x_2 = 0$, the element x_3^2 cannot be contained in $R(x_1, x_2)$; this shows that R cannot be proper. By our main theorem, the Segre product A of R with any H-Macaulay ring of altitude > 1, say the polynomial ring $S = K[y_1, y_2]$ of two variables, cannot be H-Macaulay. In fact, the elements

$$w_1 = x_1y_1, \quad w_2 = x_2y_2, \quad w_3 = (x_1 + x_2)(y_1 + y_2)$$

form a system of homogeneous parameters for the maximal homogeneous ideal $\mathfrak{L}$ of A; but this system is not distinct, for the element $x_3^2y_1y_2$ in A is not in $R(w_1, w_2)$ while the product $x_3^2y_1y_2 \cdot w_3 = x_3^2(y_1 + y_2)(y_2w_1 + y_1w_2)$ is contained in $R(w_1, w_2)$. The next example shows that even in case $m = 1$ the condition in our main theorem need not be satisfied. Let $R = K[x_1, x_2]$ be the homogeneous coordinate ring of the irreducible 0-dimensional variety defined by the equation $x_2^2 - 2x_1^2 = 0$ over a prime field K of characteristic $\neq 2$. It is clear that R is H-Macaulay with x_1 as a parameter for $\mathfrak{M}$, but R is not proper since x_2 is not contained in $R(x_1)$. Therefore, by our main theorem, the Segre product A of R with $S = K[y_1, y_2]$ is not H-Macaulay. We note that A is here the homogeneous coordinate ring $K[z_1, z_2, z_3, z_4]$ of the variety defined by the equation $z_3^2 - 2z_1^2 = 0$, $z_4^2 - 2z_2^2 = 0$, $z_1z_4 - z_2z_3 = 0$; the first two equations define four lines in the projective 3-space, while the last equation excludes two of these four lines.

THE JOHNS HOPKINS UNIVERSITY.

MAXIMALE HOLOMORPHE UND MEROMORPHE ABBILDUNGEN, II.

Von Karl Stein.*

Im zweiten Teil dieser Arbeit[1] werden die Betrachtungen über maximale Abbildungen auf den Fall meromorpher Abbildungen ausgedehnt.

§3 und §4 haben vorbereitenden Charakter. In §3 werden die Begriffe *stetige* und *holomorphe Korrespondenz* sowie damit in Zusammenhang stehende Begriffsbildungen präzisiert und einige darauf bezügliche Aussagen bewiesen. §4 handelt von allgemeinen Eigenschaften meromorpher Abbildungen, insbesondere werden Verknüpfungseigenschaften diskutiert. Der Begriff *meromorphe Abbildung* wird hier im Sinne der Definition von R. Remmert [20a, 22] benutzt.

§5 enthält die Hauptergebnisse. Es werden zunächst vorbereitende Aussagen zum Begriff der Abhängigkeit meromorpher Abbildungen (der im Falle holomorpher Abbildungen die Abhängigkeit im weiteren Sinne entspricht) zusammengestellt. Nach Präzisierung des Begriffs der *Maximalität meromorpher Abbildungen,* die als *m-Maximalität* bezeichnet wird, der *komplexen m-Basis zu einer gegebenen meromorphen Abbildung* und verwandter Begriffe werden sodann Sätze über die Existenz meromorpher Abbildungen mit Maximalitätseigenschaften bewiesen. Es ergibt sich, dass—in Analogie zu Satz I im ersten Teil der Arbeit—die Existenz einer komplexen m-Basis zu einer meromorphen Abbildung f eines irreduziblen komplexen Raumes X stets dann gesichert ist, wenn es eine analytische Menge A in X gibt, derart, dass die holomorphe Korrespondenz $f \mid A$ wenigstens eine irreduzible Komponente besitzt, die eigentlich ist und den gleichen Rang wie f hat. Diese Aussage lässt sich verschärfen: Unter der gleichen Voraussetzung wird sogleich die weitergehende Aussage (Satz III) gewonnen, dass auch in jeder "f-Klasse" $\mathfrak{F}$ meromorpher Abbildungen stets eine "$\mathfrak{F}$-maximale" Abbildung existiert (Definitionen in 5.3). Ein entsprechender Satz gilt für meromorphe

---

Received March 5, 1964.

* The author was supported by AFOSR through the European Office of Aerospace Research, U. S. Air Force Grant No. AF-EOAR-62-40.

[1] Teil I ist erschienen in Bd. 85 dieser Zeitschrift, pp. 298-315.—Berichtigung: In §1, p. 302, Zeile 29, muss es heissen "nichtsingulär" statt "irreduzibel."

Abbildungskeime. Es folgt insbesondere, dass zu jeder meromorphen Abbildung eines irreduziblen komplexen Raumes, die eine Unbestimmtheitsstelle maximaler Stufe besitzt, stets eine komplexe m-Basis existiert.

In § 6 werden Anwendungen gegeben. Mit Hilfe der Resultate von § 5 werden für Systeme meromorpher Abbildungen Aussagen gewonnen, die für den Fall meromorpher Funktionen bekannten Sätzen über Beziehungen zwischen analytischer und algebraischer Abhängigkeit entsprechen. Durch Spezialisierung ergibt sich so u. a. der Satz von W. Thimm [34, 36] über die algebraische Abhängigkeit von meromorphen Funktionen mit gemeinsamen Unbestimmtheitsstellen.

3. **Korrespondenzen.** Wir stellen eine Reihe von Erklärungen und Aussagen zusammen (vg. hierzu auch [5, 5a]).

3.1. X und Y seien Mengen. Unter einer *Korrespondenz von X in Y* wird eine Abbildung von X in die Potenzmenge $\mathfrak{P}(Y)$ von Y verstanden; an Stelle von $f\colon X \to \mathfrak{P}(Y)$ schreiben wir im folgenden stets $f\colon X \xrightarrow[k]{} Y$. Jedem Element $x \in X$ ist vermöge f also eine (leere oder nichtleere) Untermenge $f(x)$ von Y zugeordnet. Im folgenden wird oft zwischen einelementigen Mengen und ihren Elementen nicht unterschieden; dementsprechend werden Abbildungen von X in Y als spezielle Korrespondenzen aufgefasst.

Jede Korrespondenz $f\colon X \xrightarrow[k]{} Y$ ist durch eine Untermenge der Produktmenge $X \times Y$, ihren Graphen $G[f]$, charakterisiert: Es ist $G[f] = \{(x, y) : x \in X, y \in f(x)\}$, und jede Untermenge $G \subset X \times Y$ bestimmt eine Korrespondenz $f\colon X \xrightarrow[k]{} Y$ mit $G[f] = G$. Man hat zwei Abbildungen $\pi[f]\colon G[f] \to X$ und $\hat{f}\colon G[f] \to Y$, die durch die Projektionen von $X \times Y$ auf X und Y induziert werden. $\pi[f]$ heisst die Projektion von $G[f]$ in X, $\hat{f}$ die f durch Liftung nach $G[f]$ zugeordnete Abbildung.

Wir sagen, die Korrespondenz $f\colon X \xrightarrow[k]{} Y$ sei in der Korrespondenz $'f\colon X \xrightarrow[k]{} Y$ enthalten—in Zeichen $f \subset {'f}$—wenn der Graph $G[f]$ im Graphen $G['f]$ enthalten ist. f heisst leer, wenn $G[f]$ leer ist.

Ist A eine Untermenge von X, so wird unter der Bildmenge $f(A)$ vermöge der Korrespondenz $f\colon X \xrightarrow[k]{} Y$ die Menge aller $y \in Y$ mit $y \in f(x)$, $x \in A$, verstanden. Die Beschränkung von f auf A ist eine Korrespondenz von A in Y, wir bezeichnen sie mit $f \mid A$. Ist A' eine Untermenge von Y, in der die Bildmenge $f(X)$ enthalten ist, so induziert f, indem jeweils $f(x)$,

$x \in X$, als Untermenge von A' genommen wird, eine Korrespondenz f': $X \xrightarrow[k]{} A'$, die als die Beschränkung von f auf A' bezeichnet wird.

Zu jeder Korrespondenz $f: X \xrightarrow[k]{} Y$ gehört eine Umkehrkorrespondenz $f^{-1}: Y \xrightarrow[k]{} X$: Für $y \in Y$ ist $f^{-1}(y)$ als die Menge der Elemente $x \in X$ mit $y \in f(x)$ erklärt. Der Graph $G[f^{-1}] \subset Y \times X$ geht vermöge der durch $(x, y) \rightarrow (y, x)$ gegebenen Abbildung von $X \times Y$ auf $Y \times X$ aus dem Graphen $G[f]$ hervor, woraus $(f^{-1})^{-1} = f$ folgt.

Sind $f: X \xrightarrow[k]{} Y$, $f_1: X_1 \xrightarrow[k]{} Y_1$ Korrespondenzen, so ist das *kartesische Produkt* $f \times f_1: X \times X_1 \xrightarrow[k]{} Y \times Y_1$ definiert als die durch $(x, x_1) \rightarrow f(x) \times f_1(x_1) \subset Y \times Y_1$ $(x \in X, x_1 \in X_1)$ gegebene Korrespondenz. (Ist $f(x)$ oder $f_1(x_1)$ leer, so ist $f(x) \times f_1(x_1)$ die leere Untermenge von $Y \times Y_1$). Es ist $(f \times f_1)^{-1} = f^{-1} \times f_1^{-1}$. Man hat eine bijektive Abbildung τ' von $G[f] \times G[f_1] \subset X \times Y \times X_1 \times Y_1$ auf $G[f \times f_1] \subset X \times X_1 \times Y \times Y_1$, die von der durch Koordinatenpermutation bestimmten Abbildung

$$\tau: X \times Y \times X_1 \times Y_1 \rightarrow X \times X_1 \times Y \times Y_1$$

induziert wird. Im Falle $X = X_1$ wird unter der *Verbindung* (f, f_1): $X \xrightarrow[k]{} Y \times Y_1$ von f und f_1 die durch die Zuordnung $x \rightarrow f(x) \times f_1(x) \subset Y \times Y_1$ gegebene Korrespondenz verstanden. Bezeichnet D_X die Diagonale in $X \times X$, δ_X ihre Projektion auf X und ι_1 die identische Abbildung von $X_1 \times Y_1$, so ist $G[(f, f_1)] = (\delta_X \times \iota_1)(G[(f \times f_1) \mid D_X])$.

Das *Produkt* $g \circ f$ von zwei Korrespondenzen $f: X \xrightarrow[k]{} Y$, $g: Y \xrightarrow[k]{} Z$ ist definiert als diejenige Korrespondenz von X in Z, die jedem $x \in X$ die Untermenge $g(f(x)) \subset Z$ zuordnet. Für jede Untermenge $A \subset X$ ist dann $(g \circ f)(A) = g(f(A))$. Der Graph $G[g \circ f]$ ist das "Kompositum" $G[f] \circ G[g]$ von $G[f]$ und $G[g]$: Ein Paar $(x, z) \subset X \times Z$ gehört genau dann zu $G[g \circ f]$, wenn ein $y \in Y$ existiert, derart, dass $(x, y) \in G[f]$ und $(y, z) \in G[g]$ gilt. Ist D^* die Teilmenge von $X \times Y \times Y \times Z = \{(x, y, y_1, z)\}$ mit $y = y_1$ und bezeichnet β die Projektion von $X \times Y \times Y \times Z$ auf $X \times Z$, so ist $G[g \circ f] = \beta((G[f] \times G[g]) \cap D^*)$. Es ist auch $G[g \circ f] = (f^{-1}, g)(Y)$.

Es gelten die Regeln $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ und $h \circ (g \circ f) = (h \circ g) \circ f$. Im allgemeinen ist $f^{-1} \circ f$ nicht die identische Abbildung von X und $f \circ f^{-1}$ nicht die identische Abbildung von Y (ist jedoch f eine Abbildung im gewöhnlichen Sinne und surjektiv, so ist $f \circ f^{-1}$ die Identität). $(f^{-1} \circ f)(A) \supset A$ gilt für jede Teilmenge $A \subset X$ genau dann, wenn $f(x)$ für kein $x \in X$ leer oder—was dasselbe besagt—wenn f^{-1} surjektiv ist. Stets ist $f = \hat{f} \circ \pi[f]^{-1}$, aber i. a. $\hat{f} \neq f \circ \pi[f]$.

3.2. Es seien jetzt X, Y usw. topologische Räume. Wir nennen eine Korrespondenz $f: X \underset{k}{\longrightarrow} K$ *stetig*, wenn die Projektion $\pi[f]: G[f] \to X$ eine eigentliche Abbildung des Unterraums $G[f]$ von $X \times Y$ ist.[2] Es ist klar, dass diese Definition für den Spezialfall von Abbildungen $X \to Y$ den üblichen Stetigkeitsbegriff liefert. Die zu einer stetigen Abbildung $f: X \to Y$ gehörende Umkehrkorrespondenz $f^{-1}: Y \underset{k}{\longrightarrow} X$ ist genau dann stetig, wenn f eigentlich ist. Denn $f = \hat{f} \circ \pi[f]^{-1}$ ist, da $\pi[f]$ ein Homöomorphismus ist, genau dann eigentlich, wenn $\hat{f}$ es ist.

Sei $f: X \underset{k}{\longrightarrow} Y$ stetig. Ist A ein Unterraum von X, weiter A' ein Unterraum von Y, derart, dass $f(X) \subset A'$ gilt, so sind auch die Beschränkungen $f|A: A \underset{k}{\longrightarrow} Y$ und $f': X \underset{k}{\longrightarrow} A'$ von f auf A bezw. A' stetig.

Ist $f: X \underset{k}{\longrightarrow} Y$ stetig und Y Hausdorffsch, so ist der Graph $G[f]$ abgeschlossen in $X \times Y$. Denn sei $(x, y) \in X \times Y$ ein Punkt, der nicht zu $G[f]$ gehört. Dann gehört y nicht zu der kompakten Menge $f(x) \in Y$, und es gibt eine zu $f(x)$ punktfremde abgeschlossene Umgebung $\bar{V}$ von y. Die Urbildmenge $f^{-1}(\bar{V}) = \pi[f](\hat{f}^{-1}(\bar{V}))$ ist abgeschlossen in X und enthält x nicht; daher ist $U := X - f^{-1}(\bar{V})$ eine offene Umgebung von x, und es ist $U \times \bar{V} \subset X \times Y$ eine Umgebung von (x, y), die $G[f]$ nicht trifft.

PROPOSITION 3.2.1. *Die Korrespondenzen* $f: X \underset{k}{\longrightarrow} Y$, $f_1: X_1 \underset{k}{\longrightarrow} Y_1$, $f_1': X \underset{k}{\longrightarrow} Y_1$, $g: Y \underset{k}{\longrightarrow} Z$ *seien stetig. Dann sind auch*

$$f \times f_1: X \times X_1 \underset{k}{\longrightarrow} Y \times Y_1, \quad (f, f_1'): X \underset{k}{\longrightarrow} Y \times Y_1, \quad g \circ f: X \underset{k}{\longrightarrow} Z$$

stetige Korrespondenzen.

Beweis. 1) $f \times f_1$ ist stetig: Die Abbildung $\tau': G[f] \times G[f_1] \to G[f \times f_1]$ (wir verwenden hier und im folgenden die in 3.1 eingeführten Bezeichnungen) ist ein Homöomorphismus. Es ist $(\pi[f] \times \pi[f_1]) \circ \tau'^{-1} = \pi[f \times f_1]$; mit $\pi[f]$ und $\pi[f_1]$ ist $\pi[f] \times \pi[f_1]$, also auch $\pi[f \times f_1]$ eigentlich.

2) $g \circ f$ ist stetig: Sei $M \subset G[g \circ f]$ abgeschlossen. Bezeichnet

$$\beta': (G[f] \times G[g]) \cap D^* \to G[g \circ f]$$

---

[2] Über eigentliche Abbildungen siehe [5a], § 10. Eine Abbildung $X \to Y$ ist genau dann eigentlich, wenn sie stetig und abgeschlossen ist und wenn die Urbilder der Punkte von Y quasi-kompakt sind (die Urbilder quasi-kompakter Mengen in Y sind dann ebenfalls stets quasi-kompakt). Ist X Hausdorffsch und Y lokal kompakt, so ist eine stetige Abbildung $X \to Y$ dann und nur dann eigentlich, wenn die Urbilder kompakter Mengen kompakt sind.—Vgl. zu den folgenden Ausführungen auch K. Wolffhardt [38].

die durch Beschränkung von $\beta: X \times Y \times Y \times Z \rightarrow X \times Z$ bestimmte Abbildung, so ist auch $\beta'^{-1}(M) =: M^* \subset (G[f] \times G[g]) \cap D^*$ abgeschlossen. Sei M^*_1 eine abgeschlossene Menge in $G[f] \times G[g]$, sodass

$$M^*_1 \cap ((G[f] \times G[g]) \cap D^*) = M^*,$$

ferner $\iota_f: G[f] \rightarrow G[f]$ die identische Abbildung; dann ist

$$\iota_f \times \pi[g]: G[f] \times G[g] \rightarrow G[f] \times Y$$

eine eigentliche Abbildung, also ist $(\iota_f \times \pi[g])(M^*_1) =: M'_1$ abgeschlossen in $G[f] \times Y$. Es bezeichne weiter D_Y die Diagonale von $Y \times Y$, dann ist $M'_1 \cap (X \times D_Y) = (\iota_f \times \pi[g])(M^*) =: M'$ und daher M' abgeschlossen in $X \times D_Y$. Die durch die Projektion von $X \times Y \times Y$ auf $X \times Y$ induzierte Abbildung $\lambda: X \times D_Y \rightarrow X \times Y$ ist ein Homöomorphismus, also ist $\lambda(M') =: M''$ abgeschlossen in $X \times Y$. Es ist dann auch $\pi[g](M'')$ abgeschlossen in X; andererseits ist $\pi[g](M'') = \pi[g \circ f](M)$. Somit ist $\pi[g \circ f]$ abgeschlossen. Für jeden Punkt $x \in X$ ist ferner $f(x) \subset Y$ quasi-kompakt, also gilt gleiches für $g(f(x)) = g \circ f(x) \subset Z$; da $g \circ f(x)$ auf $\pi[g \circ f]^{-1}(x) \subset G[g \circ f]$ homöomorph bezogen ist, ist mithin $\pi[g \circ f]^{-1}(x)$ quasi-kompakt. $\pi[g \circ f]$ ist demnach eigentlich, und es folgt die Behauptung.

3) Die Abbildung $\delta_X: D_X \rightarrow X$ ist ein Homöomorphismus; es ist $(f, f_1') = (f \times f_1' \mid D_X) \circ \delta_X^{-1}$. Auf Grund von 1) and 2) ist (f, f_1') somit stetig.

3.3. Eine Korrespondenz $f: X \xrightarrow[k]{} Y$ heisse eigentlich, wenn f und die Umkehrkorrespondenz $f^{-1}: Y \xrightarrow[k]{} X$ stetig sind[3] (f^{-1} ist genau dann stetig, wenn $\hat{f}: G[f] \rightarrow Y$ eine eigentliche Abbildung ist). Sind die Räume X und Y lokal kompakt, so ist eine stetige Korrespondenz $f: X \xrightarrow[k]{} Y$ dann und nur dann eigentlich, wenn für jede kompakte Menge $K_Y \subset Y$ die Menge $f^{-1}(K_Y)$ in X kompakt ist.

PROPOSITION 3.3.1. *Es seien* $f: X \xrightarrow[k]{} Y$, $f_1: X_1 \xrightarrow[k]{} Y_1$, $g: Y \xrightarrow[k]{} Z$ *eigentliche Korrespondenzen. Dann sind auch* $f \times f_1$ *und* $g \circ f$ *eigentliche Korrespondenzen.*

Beweis. Es ist $(f \times f_1)^{-1} = f^{-1} \times f_1^{-1}$ und $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$; nach Proposition 3.2.1 sind mit f, f^{-1}, f_1, f_1^{-1}, g, g^{-1} auch $f \times f_1$, $f^{-1} \times f_1^{-1}$, $g \circ f$, $f^{-1} \circ g^{-1}$ stetig.

---

[3] Vgl. [5a], p. 172, Ex. 10 ff., wo der Begriff *eigentlich* (*propre*) für Korrespondenzen weiter gefasst ist. Die oben benutzte Fassung führt für den Sonderfall von Abbildungen $X \rightarrow Y$ zum üblichen Begriff der Eigentlichkeit zurück.—Einige der obigen Aussagen ergeben sich auch aus den Exercises in [5a].

Folgerung: Ist A eine abgeschlossene Menge in X, so ist mit $f: X \xrightarrow[k]{} Y$ auch die Beschränkung $f \mid A$ von f auf den Unterraum $A \subset X$ eigentlich.

Denn die Injektion $\epsilon: A \to X$ ist eine eigentliche Abbildung, und es ist $f \mid A = f \circ \epsilon$.

PROPOSITION 3.3.2. *Es seien $f: X \xrightarrow[k]{} Y$, $f_1': X \xrightarrow[k]{} Y_1$, $g: Y \xrightarrow[k]{} Z$ stetige Korrespondenzen. Die Räume X, Y, Y_1, Z seien lokal kompakt. Dann gilt:*

(I) *Ist f eigentlich, so auch (f, f_1') und (f_1', f).*

(II) *Ist $g \circ f$ eigentlich und g^{-1} surjektiv, so ist f eigentlich.*

(III) *Ist $g \circ f$ eigentlich und f surjektiv, so ist g eigentlich.*

Beweis. Es seien K_Y, K_{Y_1}, K_Z kompakte Mengen in Y bezw. Y_1 bezw. Z.

(I) Nach Proposition 3.2.1 sind (f, f_1') und (f_1', f) stetig. Es ist weiter $f^{-1}(K_Y)$ kompakt, und $f_1'^{-1}(K_{Y_1})$ ist abgeschlossen in X, daher ist $f^{-1}(K_Y) \cap f_1'^{-1}(K_{Y_1}) = (f, f_1')^{-1}(K_Y \times K_{Y_1})$ kompakt. Da jede kompakte Teilmenge von $Y \times Y_1$ in einer kompakten Teilmenge der Form $K_Y \times K_{Y_1}$ enthalten ist, folgt die Eigentlichkeit von (f, f_1'). Entsprechend ergibt sich, dass (f_1', f) eigentlich ist.

(II) Die Menge $f^{-1}(K_Y) = \pi[f](\hat{f}^{-1}(K_Y))$ ist abgeschlossen in X. Wegen der Surjektivität von g^{-1} gilt $K_Y \subset g^{-1}(g(K_Y))$, also ist

$$f^{-1}(K_Y) \subset f^{-1}(g^{-1}(g(K_Y))) = (g \circ f)^{-1}(g(K_Y));$$

mit $g(K_Y)$ ist aber $(g \circ f)^{-1}(g(K_Y))$ kompakt, da $g \circ f$ als eigentlich vorausgesetzt ist. $f^{-1}(K_Y)$ ist demnach kompakt und mithin f eigentlich.

(III) Für jede Teilmenge $A_Y \subset Y$ gilt $A_Y \subset (f \circ f^{-1})(A_Y)$, da f surjektiv ist; also ist $g^{-1}(K_Z) \subset (f \circ f^{-1})(g^{-1}(K_Z)) = f((g \circ f)^{-1}(K_Z))$. Es ist $g^{-1}(K_Z) = \pi[g](g^{-1}(K_Z))$ in X abgeschlossen; andererseits ist $f((g \circ f)^{-1}(K_Z))$ wegen der Eigentlichkeit von $g \circ f$ kompakt. Es folgt, dass $g^{-1}(K_Z)$ kompakt und somit g eigentlich ist.

3.4. Im folgenden seien X, Y usw. komplexe Räume.[4] Eine stetige Korrespondenz $f: X \xrightarrow[k]{} Y$ heisse *holomorph*, wenn der Graph $G[f]$ eine analytische Menge im komplexen Raum $X \times Y$ ist.

---

[4] Unter einem *komplexen Raum* wird in dieser Arbeit stets ein komplexer Raum im Sinne von J. P. Serre [27] (*reduzierter komplexer Raum* in der Terminologie von H. Grauert [10]) verstanden; alle vorkommenden komplexen Räume werden als nichtleer und endlichdimensional vorausgesetzt.

Sei $f: X \xrightarrow[k]{} Y$ nichtleer und holomorph, dann sind $\pi[f]: G[f] \to X$ und $\hat{f}: G[f] \to Y$ holomorphe Abbildungen. Unter dem Rang $r(f)$ von f wird der Rang $r(\hat{f})$ verstanden. Ist A eine nichtleere lokal-analytische Menge in X, weiter A' eine lokal-analytische Menge in Y, derart, dass $f(X) \subset A'$ gilt, so sind auch die Beschränkungen $f \mid A: A \xrightarrow[k]{} Y$ und $f': X \xrightarrow[k]{} A'$ von f auf A bezw. A' holomorph. Ist A'' eine analytische Menge in Y, so ist $f^{-1}(A'')$ eine analytische Menge in X. Denn $\hat{f}^{-1}(A'')$ ist analytisch in $G[f]$; da $\pi[f]$ eigentlich ist, folgt nach R. Remmert [22] die Analytizität von $\pi[f](\hat{f}^{-1}(A'')) = f^{-1}(A'')$ in X. Ferner: Für jeden Punkt $x \in X$ ist $f(x)$ eine (leere oder nichtleere) kompakte analytische Menge in Y, denn $\pi[f]^{-1}(x) \subset G[f]$ ist eine kompakte analytische Menge. Ist f auch eigentlich, so ist $f(X) = \hat{f}(G[f])$ eine $r(f)$-dimensionale analytische Menge in Y.

PROPOSITION 3.4.1. *Die Korrespondenzen* $f: X \xrightarrow[k]{} Y$, $f_1: X_1 \xrightarrow[k]{} Y_1$, $f_1': X \xrightarrow[k]{} Y_1$, $g: Y \xrightarrow[k]{} Z$ *seien holomorph. Dann sind auch*

$$f \times f_1: X \times X_1 \xrightarrow[k]{} Y \times Y_1, \quad (f, f_1'): X \xrightarrow[k]{} Y \times Y_1, \quad g \circ f: X \xrightarrow[k]{} Z$$

holomorphe Korrespondenzen.

Zum Beweis genügt es wegen Proposition 3.2.1 zu zeigen, dass die Graphen der Korrespondenzen $f \times f_1$, (f, f_1') und $g \circ f$ analytische Mengen sind. Für $G[f \times f_1]$ und $G[(f, f_1')]$ ist dies klar. Weiter ist (mit den Bezeichnungen von 3.1) $G[g \circ f] = \beta((G[f] \times G[g]) \cap D^*)$. Es ist $(G[f] \times G[g]) \cap D^*$ eine analytische Menge in $X \times Y \times Y \times Z$. Die holomorphe Abbildung $\beta \mid (G[f] \times G[g]) \cap D^* =: \beta^*: (G[f] \times G[g]) \cap D^* \to X \times Z$ ist ferner eigentlich. Denn ist $K \subset X \times Z$ kompakt, so ist zunächst $\beta^{*-1}(K) \subset (G[f] \times G[g]) \cap D^*$ abgeschlossen; bezeichnet σ_X bezw. σ_Z die Projektion von $X \times Z$ auf X bezw. Z, so ist andererseits $\beta^{*-1}(K) = \beta^{-1}(K) \cap (G[f] \times G[g]) \cap D^*$ in der kompakten Menge

$$\pi[f]^{-1}(\sigma_X(K)) \times f(\sigma_X(K)) \times \sigma_Z(K) \subset G[f] \times Y \times Z$$

enthalten. Es folgt, dass $\beta^{*-1}(K)$ kompakt und somit β^* eigentlich ist; also ist $\beta^*((G[f] \times G[g]) \cap D^*) = G[g \circ f]$ nach [22] eine in $X \times Z$ analytische Menge.

Eine nichtleere holomorphe Korrespondenz $f: X \xrightarrow[k]{} Y$ heisse *irreduzibel*, wenn ihr Graph $G[f]$ eine irreduzible analytische Menge in $X \times Y$ ist; anderenfalls heisst f *reduzibel*. Ist $G[f] = \bigcup_i G^{(i)}$ die Zerlegung von $G[f]$

in irreduzible Komponenten, so heissen die Korrespondenzen $f_i\colon X \underset{k}{\longrightarrow} Y$ mit $G[f_i] := G^{(i)}$ die *irreduziblen Komponenten* von f; wir schreiben $f = \bigcup_i f_i$. Es ist klar, dass jede irreduzible Komponente f_i von f wieder holomorph ist; ferner: Ist f eigentlich, so gilt dies auch für jedes f_i, allgemeiner für jede in f enthaltene holomorphe Korrespondenz.

Wir benötigen im folgenden oft

PROPOSITION 3.4.2. *Es seien* $f_j\colon X \underset{k}{\longrightarrow} Y$ $(j = 1, 2)$ *irreduzible holomorphe Korrespondenzen; es existiere eine offene Menge* U *in* X, *derart, dass die Beschränkungen* $f_j \mid U\colon U \underset{k}{\longrightarrow} Y$ *nichtleer sind und übereinstimmen. Dann ist* $f_1 = f_2$.

Zum Beweis ist nur zu beachten, dass im Durchschnitt der Graphen $G[f_1]$ und $G[f_2]$, die irreduzible analytische Mengen in $X \times Y$ sind, die in $G[f_1]$ und $G[f_2]$ offene nichtleere Menge $G[f_j \mid U]$ enthalten ist; daraus folgt $G[f_1] = G[f_2]$ und $f_1 = f_2$.

4. Meromorphe Abbildungen.

4.1. Sei $f\colon X \underset{k}{\longrightarrow} Y$ eine holomorphe Korrespondenz. Ein Punkt $x_0 \in X$ heisse ein *ausgezeichneter Punkt* inbezug auf f, wenn es eine Umgebung U von x_0 gibt, derart, dass für jeden Punkt $x \in U$ die Menge $f(x)$ einpunktig, dass also $f \mid U\colon U \underset{k}{\longrightarrow} Y$ eine Abbildung im gewöhnlichen Sinne ist. Eine nichtleere Menge in X, die aus lauter ausgezeichneten Punkten inbezug auf f besteht, nennen wir *ausgezeichnet* inbezug auf f. Ist U eine offene Menge dieser Art, so ist die holomorphe Korrespondenz $\pi[f \mid U]^{-1}\colon U \underset{k}{\longrightarrow} G[f \mid U]$ eine stetige Abbildung von U in $G[f \mid U]$, und diese Abbildung ist in jedem gewöhnlichen Punkte von U holomorph. Demnach hat auch f in jedem gewöhnlichen Punkt von U den Charakter einer holomorphen Abbildung. Da die nichtgewöhnlichen Punkte des komplexen Raumes U eine in U nirgendsdichte analytische Menge bilden, folgt, dass es in U eine dichte offene Menge U' gibt, sodass $f \mid U'$ eine holomorphe Abbildung ist.

Eine nichtleere holomorphe Korrespondenz $f\colon X \underset{k}{\longrightarrow} Y$ heisst eine *meromorphe Abbildung* von X in Y, wenn folgendes zutrifft: Ist der komplexe Raum X irreduzibel, so gilt

1) f ist irreduzibel,

2) es gibt in X wenigstens einen ausgezeichneten Punkt inbezug auf f; ist X reduzibel und $X = \bigcup_i X^{(i)}$, $i \in I$, die Zerlegung von X in irreduzible Komponenten, so gibt es holomorphe Korrespondenzen $f_i \colon X \underset{k}{\longrightarrow} Y$, derart, dass $f = \bigcup_i f_i$ gilt und dass jeweils $f_i \mid X^{(i)}$ eine meromorphe Abbildung und $f_i \mid X - X^{(i)}$ leer ist.[5] Ist f eine meromorphe Abbildung von X in Y, so schreiben wir auch $f \colon X \underset{m}{\longrightarrow} Y$. Besitzt ein Punkt $x \in X$ eine offene Umgebung U, derart, dass $f \mid U$ eine holomorphe Abbildung ist, so nennen wir x *nichtsingulär* inbezug auf die meromorphe Abbildung f; anderenfalls heisst x *eine ausserwesentliche Singularität* von f. Jeder inbezug auf $f \colon X \underset{m}{\longrightarrow} Y$ nichtsinguläre Punkt ist ausgezeichnet inbezug auf die holomorphe Korrespondenz f; umgekehrt gibt es in jeder inbezug auf die holomorphe Korrespondenz f ausgezeichneten offenen Menge eine offene dichte Untermenge, deren Punkte nichtsingulär inbezug auf die meromorphe Abbildung f sind. Die Menge der ausserwesentlichen Singularitäten von $f \colon X \underset{m}{\longrightarrow} Y$ werde mit $S(f)$ bezeichnet. Zu $S(f)$ gehören insbesondere diejenigen Punkte $x \in X$, für welche $f(x)$ nicht einpunktig ist; solche Punkte heissen *Unbestimmtheitsstellen* von f.

Meromorphe Abbildungen hängen eng mit eigentlichen Modifikationen komplexer Räume zusammen. Wir erinnern zunächst an diesen Begriff (vgl. hierzu [3, 11, 16, 17, 31, 32]). Sei $\tau \colon {}'X \to X$ eine eigentliche holomorphe Abbildung; es existiere eine offene Menge U in X mit folgenden Eigenschaften: 1) U hat mit jeder irreduziblen Komponente von X gemeinsame Punkte und $\tau^{-1}(U)$ hat mit jeder irreduziblen Komponente von ${}'X$ gemeinsame Punkte, 2) die Korrespondenz $\tau^{-1} \mid U \colon U \underset{k}{\longrightarrow} X$ ist eine holomorphe Abbildung. Dann heisst τ eine *eigentliche Modifikationsabbildung* und das Tripel $({}'X, \tau, X)$ eine *eigentliche Modifikation* von X. Ein Punkt von X heisst *singulär* inbezug auf $({}'X, \tau, X)$, wenn er keine offene Umgebung V besitzt, derart, dass $\tau^{-1} \mid V \colon V \underset{k}{\longrightarrow} {}'X$ eine holomorphe Abbildung ist.

Eine eigentliche Modifikationsabbildung $\tau \colon {}'X \to X$ ist insbesondere surjektiv. Ist ${}'X = \bigcup_i {}'X^{(i)}$, $i \in I$, die Zerlegung von ${}'X$ in irreduzible Komponenten, so entspricht dieser Zerlegung die von X in irreduzible Komponenten: Es ist jeweils $\tau({}'X^{(i)}) =: X^{(i)}$ eine irreduzible Komponente von X, für $i \neq j$ sind $X^{(i)}$ und $X^{(j)}$ verschieden, und es gilt $X = \bigcup_i X^{(i)}$. Die durch Beschränkung von τ bestimmten Abbildungen $\tau_i\colon {}'X^{(i)}\to X^{(i)}$ sind ebenfalls eigentliche Modifikationsabbildungen, und es ist jeweils $r(\tau_i)=\dim{}'X^{(i)}=\dim X^{(i)}$.

---

[5] Der hier benutzte Begriff der meromorphen Abbildung eines komplexen Raumes stammt von R. Remmert [20a, 21]. Ein anderer Begriff der meromorphen Abbildung wurde von W. Stoll [31] (vgl. auch [1, 33] eingeführt. Über meromorphe Abbildungen und ihre Grundeigenschaften siehe ferner [16, 17, 18a].

Eine holomorphe Korrespondenz $f\colon X\xrightarrow[k]{}Y$ ist nun offenbar genau dann eine meromorphe Abbildung, wenn die holomorphe Abbildung $\pi[f]\colon G[f]\to X$ eine eigentliche Modifikationsabbildung ist. Hat $\pi[f]$ diese Eigenschaft, so ist jeder inbezug auf die eigentliche Modifikation $(G[f],\pi[f],X)$ singuläre Punkt von X eine ausserwesentlich Singularität der meromorphen Abbildung f und umgekehrt.—Wir notieren ferner, dass die zu einer eigentlichen Modifikationsabbildung $\tau\colon {}'X\to X$ gehörige Umkehrkorrespondenz $\tau^{-1}\colon X\xrightarrow[k]{}{}'X$ stets eine meromorphe Abbildung ist.

Es gilt:

PROPOSITION 4.1.1. *Die Menge der singulären Punkte inbezug auf eine eigentliche Modifikation* $('X,\tau,X)$ *ist eine nirgendsdichte analytische Menge in* X.

Hieraus folgt unmittelbar

PROPOSITION 4.1.2. *Die Menge* $S(f)$ *der ausserwesentlichen Singularitäten einer meromorphen Abbildung* $f\colon X\xrightarrow[m]{}Y$ *ist eine nirgendsdichte analytische Menge in* X.

Wir stützen den Beweis von Proposition 4.1.1 auf verschiedene Hilfsaussagen.[6]

Sei $\phi\colon X\to Y$ eine holomorphe Abbildung. ϕ heisst im Punkt $x\in X$ *regulär,* wenn der durch ϕ induzierte Homomorphismus des Ringes der in $y=\phi(x)\in Y$ holomorphen Funktionskeime in den Ring der in $x\in X$ holomorphen Funktionskeime surjektiv ist. Es gilt

(I) *Die Menge der Punkte, in denen eine holomorphe Abbildung* $\phi\colon X\to Y$ *nicht regulär ist, ist eine analytische Menge in* X.

(II) *Die holomorphe Abbildung* $\phi\colon X\to Y$ *ist im Punkte* $x\in X$ *genau dann regulär, wenn es eine offene Umgebung* $V=V(x)$ *und eine lokalanalytische Menge* $A^*\subset Y$ *gibt, derart, dass* $\phi(V)=A^*$ *und die durch Beschränkung von* ϕ *bestimmte Abbildung* $V\to A^*$ *biholomorph ist.*

Da beide Aussagen lokaler Natur sind, genügt es zu ihrem Beweis, die

---

[6] Auf die Möglichkeit, den Beweis wie im folgenden zu führen, wurde ich von H. Grauert hingewiesen. Vgl. auch [10a], § 1.—Die unten benutzte Aussage (II) ergibt sich unmittelbar auch aus der Aussage (1) von A. Andreotti in [10a], § 1. Vgl. ferner [18b].

Betrachtung auf je eine Umgebung eines Punktes von X und seines Bildpunktes in Y zu beschränken. Es darf daher angenommen werden, dass X als eine analytische Menge in einem Gebiet D_0 des $\boldsymbol{C}^n(\xi_1,\cdots,\xi_n)=\boldsymbol{C}^n(\xi)$ und Y als eine analytische Menge in einem Gebiet D_1 des $\boldsymbol{C}^l(\eta_1,\cdots,\eta_l)=\boldsymbol{C}^l(\eta)$ gegeben ist. Ist dann $x_0\in X$ irgendein Punkt, so lässt sich ein Gebiet $D\subset D_0$, das x_0 enthält, mit der folgenden Eigenschaft wählen: Es gibt eine holomorphe Abbildung $F\colon D\to D_1$, derart, dass $F(D\cap X)\subset Y$ ist, und dass die durch Beschränkung von F bestimmte Abbildung $D\cap X\to Y$ mit $\phi\mid D\cap X$ übereinstimmt.—F wird durch

$$\xi\to\eta=(\eta_1,\cdots,\eta_l)=(f_1(\xi),\cdots,f_l(\xi))$$

beschrieben, wo $\xi\in D$, $\eta\in D_1$ ist und $f_1(\xi),\cdots,f_l(\xi)$ in D holomorphe Funktionen sind. Es sei noch $A:=D\cap X$ und $F':=F\mid A\colon A\to D_1$ gesetzt.

Wir benutzen weiter die folgende Terminologie: Ist Z ein komplexer Raum, Z' eine analytische Menge in Z, so sei $\mathfrak{O}(Z)$ die Strukturgarbe von Z und $\mathfrak{J}(Z')\subset\mathfrak{O}(Z)$ die Idealgarbe von Z'. Für einen Punkt $z\in Z$ wird der Halm über z einer Garbe über Z durch einen unteren Index z gekennzeichnet. Das maximale Ideal im Ringe $\mathfrak{O}_z(Z)$ werde mit $\mathfrak{m}_z(Z)$ bezeichnet; für $z\in Z'$ sei ferner $\mu_z[Z']\colon\mathfrak{m}_z(Z)\to\mathfrak{m}_z(Z')$ der natürliche Homomorphismus.

Ist nun a irgendein Punkt von A und $b:=\phi(a)=F'(a)\in Y\subset D_1$, so hat man durch ϕ, F, F' induzierte Homomorphismen

$$\begin{gathered}{}^*\phi_a\colon\mathfrak{O}_b(Y)\to\mathfrak{O}_a(A),\quad {}^*F_a\colon\mathfrak{O}_b(D_1)\to\mathfrak{O}_a(D),\\ {}^*F'_a\colon\mathfrak{O}_b(D_1)\to\mathfrak{O}_a(A)\ \text{sowie}\ \nu_a\colon\mathfrak{m}_b(Y)\to\mathfrak{m}_a(A),\\ \nu'_a\colon\mathfrak{m}_b(D_1)\to\mathfrak{m}_a(D)\ \text{und}\ \nu''_a\colon\mathfrak{m}_b(D_1)\to\mathfrak{m}_a(A).\end{gathered}$$

${}^*\phi_a$ ist genau dann surjektiv, wenn ${}^*F'_a$ es ist, d.h. ϕ ist im Punkte a genau dann regulär, wenn dies für F' in a zutrifft. Diese Eigenschaft hat F' andererseits genau dann, wenn ν''_a surjektiv ist, und dies tritt genau dann ein, wenn $\mathfrak{m}_a(D)=\mathfrak{J}(A)+\nu'_a(\mathfrak{m}_b(D_1))$ gilt, denn man hat $\nu''_a=\mu_a[A]\circ\nu'_a$, wo $\mu_a[A]$ surjektiv und $\operatorname{Kern}(\mu_a[A])=\mathfrak{J}_a(A)$ ist.

Es werde jetzt eine offene Umgebung $U(a)\subset D$ und ein System von in $U(a)$ holomorphen Funktionen $g_1,\cdots,g_s$ so gewählt, dass für jeden Punkt $u=(u_1,\cdots,u_n)\in U(a)\cap A$ das System der von den $g_1,\cdots,g_s$ repräsentierten Funktionskeime $g_{1,u},\cdots,g_{s,u}$ das Ideal $\mathfrak{J}_u(A)\subset\mathfrak{O}_u(D)$ erzeugt; eine solche Wahl ist wegen der Kohärenz der Garbe $\mathfrak{J}(A)\subset\mathfrak{O}(D)$ möglich. Die Elemente des Ringes $\nu'_u(\mathfrak{m}_{\phi(u)}(D_1))\subset\mathfrak{O}_u(D)$ werden repräsentiert durch die in u verschwindenden konvergenten Potenzreihen in den Funktionen $f^{(u)}{}_j(\xi):=f_j(\xi)-f_j(u)$. Es ist nun

$$\mathfrak{m}_u(D)=\mathfrak{J}_u(A)+\nu'_u(\mathfrak{m}_{\phi(u)}(D_1))$$

—also F' in u regulär—genau dann, wenn die Funktionalmatrix $\mathfrak{M}^{(u)}(\xi)$ des Systems der Funktionen $g_1(\xi), \cdots, g_s(\xi)$, $f^{(u)}{}_1(\xi), \cdots, f^{(u)}{}_l(\xi)$ im Punkte $\xi = u$ den Rang n hat. Denn ist der Rang von $\mathfrak{M}^{(u)}(\xi)$ gleich n, so lassen sich gewisse der $g_\sigma(\xi)$, $f^{(u)}{}_\lambda(\xi)$ als lokale Koordinaten in u einführen; jede in u holomorphe Funktion lässt sich dort also in eine Potenzreihe in den $g_\sigma(\xi)$, $f^{(u)}{}_\lambda(\xi)$ entwickeln, woraus unmittelbar

$$\mathfrak{m}_u(D) = \mathfrak{J}_u(A) + \nu'_u(\mathfrak{m}_{\phi(u)}(D_1))$$

folgt. Ist andererseits der Rang von $\mathfrak{M}^{(u)}(\xi)$ kleiner als n, so lassen sich sicher nicht alle lokalen Koordinaten $\xi^{(u)}{}_j := \xi_j - u_j$ $(j = 1, \cdots, n)$ in der Gestalt $\xi^{(u)}{}_j = \alpha_j(\xi) + \beta_j(\xi)$ darstellen, wo die durch die α_j bezw. β_j repräsentierten Keime $\alpha_{j,u}$ bezw. $\beta_{j,u}$ zu $\mathfrak{J}_u(A)$ bezw. $\nu'_u(\mathfrak{m}_{\phi(u)}(D_1))$ gehören; also ist $\mathfrak{m}_u(D) \neq \mathfrak{J}_u(A) + \nu'_u(\mathfrak{m}_{\phi(u)}(D_1))$.—Für jeden Punkt $\xi \in U(a)$ stimmt der Rang von $\mathfrak{M}^{(u)}(\xi)$ mit dem Rang der Funktionalmatrix $\mathfrak{M}(\xi)$ des Systems $g_1(\xi), \cdots, g_s(\xi)$, $f_1(\xi), \cdots, f_l(\xi)$ überein. Die Punkte $u \in U(a) \cap A$, in denen F' nicht regulär ist, sind demnach dadurch charakterisiert, dass in ihnen der Rang der Matrix $\mathfrak{M}(\xi)$ kleiner als n ist; diese Punkte bilden mithin eine analytische Menge in $U(a) \cap A$. Hieraus folgt die Aussage (I).

Es bezeichne weiter $H \colon U(a) \to \boldsymbol{C}^{s+l}(\zeta_1, \cdots, \zeta_s, \eta_1, \cdots, \eta_l) = \boldsymbol{C}^{s+l}(\zeta, \eta)$ die durch die Zuordnung $\xi \to (\zeta_1, \cdots, \zeta_s, \eta_1, \cdots, \eta_l) = (g_1(\xi), \cdots, g_s(\xi), f_1(\xi), \cdots, f_l(\xi))$ $(\xi \in U(a))$ bestimmte holomorphe Abbildung. Ist F' in a regulär, so ergibt sich aus der obigen Betrachtung, dass eine offene Umgebung $U_1(a) \subset U(a)$ mit der folgenden Eigenschaft existiert: $H(U_1(a))$ ist eine singularitätenfreie analytische Menge in einer offenen Umgebung von $H(a)$ im $\boldsymbol{C}^{s+l}(\zeta, \eta)$, derart, dass die durch Beschränkung von H bestimmte Abbildung $'H \colon U_1(a) \to H(U_1(a))$ biholomorph ist. Die Menge $U_1(a) \cap A$ wird daher vermöge $'H$ bijektiv auf eine analytische Untermenge $'A$ von $H(U_1(a))$ abgebildet. $'A$ liegt in dem durch $\zeta_1 = \cdots = \zeta_s = 0$ gegebenen Unterraum L^l des $\boldsymbol{C}^{s+l}(\zeta, \eta)$. Vermöge der durch $(\zeta, \eta) \to \eta$ gegebenen Projektion $\boldsymbol{C}^{s+l}(\zeta, \eta) \to \boldsymbol{C}^l(\eta)$ wird $'A$ bijektiv auf eine lokal-analytische Menge A^* in D_1 bezogen; es ist dann $F'(U_1(a) \cap A) = A^*$, und die durch Beschränkung von F' bestimmte Abbildung $F'^* \colon U_1(a) \cap A \to A^*$ ist biholomorph.—Gibt es umgekehrt zu F' ein $U_1(a)$ und ein A^* mit dieser Eigenschaft, so ist klar, dass F' in a regulär ist. Somit folgt die Aussage (II).

Wir benötigen weiter

(III) *Sei $\phi \colon X \to Y$ eine eigentliche holomorphe Abbildung, ferner N die Menge derjenigen Punkte x von X, für welche die Faser $\phi^{-1}(\phi(x))$ mehr als einen Punkt enthält. Dann ist die abgeschlossene Hülle $\bar{N}$ von N in X eine analytische Menge.*

Wir beweisen dies mittels vollständiger Induktion nach der Dimension des komplexen Raumes X.[7]—Die Behauptung ist richtig für $\dim X = 0$. Angenommen, sie sei bewiesen für Dimensionen bis $n-1$ $(n > 0)$. Sei jetzt $\dim X = n$. Ist $X = \bigcup_{\iota} X^{(\iota)}$ $(\iota \in I)$ die Zerlegung von X in irreduzible Komponenten, so werde mit ${}^{(\iota)}X$ jeweils die Vereinigung der von $X^{(\iota)}$ verschiedenen irreduziblen Komponenten von X bezeichnet. Ein $X^{(\iota)}$ ist jedenfalls dann in $\tilde{N}$ enthalten, wenn $\phi(X^{(\iota)}) \subset \phi({}^{(\iota)}X)$ oder wenn $r(\phi \mid X^{(\iota)}) < \dim X^{(\iota)}$. Sei $\phi(X^{(\iota)}) \not\subset \phi({}^{(\iota)}X)$ und $r(\phi \mid X^{(\iota)}) = \dim X^{(\iota)}$. Dann ist $\phi(X^{(\iota)}) =: X^{(\iota)*}$ eine irreduzible analytische Menge in Y mit $\dim X^{(\iota)*} = \dim X^{(\iota)}$. Es sei $\phi_\iota\colon X^{(\iota)} \to X^{(\iota)*}$ die durch Beschränkung von ϕ bestimmte eigentliche holomorphe Abbildung, ferner K_ι die Menge der Punkte von $X^{(\iota)}$, in denen ϕ_ι nicht regulär ist. Nach (I) ist K_ι eine analytische Menge in $X^{(\iota)}$, und diese ist wegen $r(\phi_\iota) = \dim X^{(\iota)}$ von $X^{(\iota)}$ verschieden. Bezeichnet L_ι^* die Menge der nichtgewöhnlichen Punkte von $X^{(\iota)*}$, und wird

$$M_\iota^* := L_\iota^* \cup \phi_\iota(K_\iota), \qquad M_\iota := \phi_\iota^{-1}(M_\iota^*)$$

gesetzt, so sind dann auch die analytischen Mengen M_ι^*, M_ι von $X^{(\iota)*}$ bezw. $X^{(\iota)}$ verschieden. Sei $\phi_\iota'\colon X^{(\iota)} - M_\iota \to X^{(\iota)*} - M_\iota^*$ die durch Beschränkung von ϕ_ι bestimmte eigentliche surjektive und wegen (II) lokal biholomorphe Abbildung, weiter b_ι die Blätterzahl der analytischen Überlagerung

$$(X^{(\iota)} - M_\iota, \phi_\iota', X^{(\iota)*} - M_\iota^*).$$

Ist $b_\iota > 1$, so gehört sicher $X^{(\iota)}$ zu $\tilde{N}$. Ist aber $b_\iota = 1$, so ist ϕ_ι' eine biholomorphe Abbildung, insbesondere ist für $x \in X^{(\iota)} - M_\iota$ die Faser

$$\phi_\iota'^{-1}(\phi_\iota'(x)) = \phi_\iota^{-1}(\phi_\iota(x))$$

stets einpunktig. Die in $X^{(\iota)}$ analytische Menge $(\phi \mid X^{(\iota)})^{-1}(\phi({}^{(\iota)}X))$ ist von $X^{(\iota)}$ verschieden, daher ist auch $X_0^{(\iota)} := M_\iota \cup (\phi \mid X^{(\iota)})^{-1}(\phi({}^{(\iota)}X))$ eine von $X^{(\iota)}$ verschiedene analytische Menge in $X^{(\iota)}$. Offenbar ist jetzt $X^{(\iota)} \cap \tilde{N}$ in $X_0^{(\iota)}$ enthalten.

Es hat sich ergeben, dass für beliebiges $\iota \in I$ entweder $X^{(\iota)} \cap \tilde{N} = X^{(\iota)}$ ist oder dass es in $X^{(\iota)}$ eine von $X^{(\iota)}$ verschiedene analytische Menge $X_0^{(\iota)}$ gibt, derart, dass $X^{(\iota)} \cap \tilde{N}$ in $X_0^{(\iota)}$ enthalten ist. Sei ${}_0X$ die Vereinigung der $X^{(\iota)}$ mit $X^{(\iota)} \cap \tilde{N} = X^{(\iota)}$ und X_0 die Vereinigung der in den übrigen $X^{(\iota)}$ enthaltenen $X_0^{(\iota)}$. Da ${}_0X$ und X_0 abgeschlossene komplexe Unterräume von X sind, sind die holomorphen Abbildungen $\phi \mid {}_0X$ und $\phi \mid X_0$ eigentlich;

---

[7] Wie früher festgelegt, werden alle vorkommenden komplexen Räume als endlichdimensional vorausgesetzt. Für beliebige komplexe Räume bleibt die Aussage offenbar gültig.

ferner ist $\dim X_0 < \dim X = n$. Bezeichnet N_0 die Menge derjenigen Punkte x_0 von X_0, für welche die Faser $(\phi \mid X_0)^{-1}((\phi \mid X_0)(x_0))$ nicht einpunktig ist, so ist nach Induktionsvoraussetzung die abgeschlossene Hülle $\bar{N}_0$ von N_0 eine analytische Menge in X_0. Es ist nun

$$\bar{N} = ({}_0X \cap \bar{N}) \cup (X_0 \cap \bar{N}) = {}_0X \cup \bar{N}_0 \cup (\phi \mid X_0)^{-1}(\phi({}_0X));$$

da ${}_0X$, $\bar{N}_0$ und $(\phi \mid X_0)^{-1}(\phi({}_0X))$ analytische Mengen in X sind, ist also auch $\bar{N}$ eine analytische Menge in X.

Nun zum Beweise von Proposition 4.1.1.—Sei $'N_1$ die Menge der Punkte von $'X$, in denen $\tau: {'X} \to X$ nicht regulär ist, ferner $'N_2$ die Menge derjenigen Punkte $'x \in {'X}$, für welche die Faser $\tau^{-1}(\tau('x))$ nicht einpunktig ist. $'N_1$ und die abgeschlossene Hülle $'\bar{N}_2$ von $'N_2$ sind nach (I) bezw. (III) analytische Mengen in $'X$; also ist $'N := {'N_1} \cup {'\bar{N}_2}$ analytisch in $'X$, und $'N$ ist in $'X$ nirgendsdicht, da τ eine eigentliche Modifikationsabbildung ist. Es ist dann auch $N := \tau('N)$ eine in X nirgendsdichte analytische Menge. Jeder Punkt $x \in N$ ist singulärer Punkt inbezug auf die eigentliche Modifikation $('X, \tau, X)$: Wäre x gewöhnlicher Punkt inbezug auf $('X, \tau, X)$, so gäbe es eine offene Umgebung $V = V(x)$, sodass $\tau^{-1} \mid V;\ V \xrightarrow[k]{} {'X}$ eine holomorphe Abbildung ist; dann aber könnte kein Punkt von $\tau^{-1}(V)$ zu $'N_1$ oder $'N_2$ gehören. Ist andererseits x_0 ein Punkt von $X - N$, so gibt es genau einen Punkt $'x_0 \in {'X}$ mit $\tau('x_0) = x_0$. Da $'x_0 \in {'X} - N$, ist τ in $'x_0$ regulär. Nach (II) gibt es mithin eine offene Umgebung $'V = {'V}('x_0)$ und eine lokal-analytische Menge $A \subset X$, derart, dass $\tau('V) = A$ und die durch Beschränkung von τ bestimmte Abbildung $'\tau: {'V} \to A$ biholomorph ist. A muss eine Umgebung von x_0 in X sein. Denn sonst gäbe es eine Folge von Punkten $x_\nu \in X - A$ $(\nu = 1, 2, \cdots)$, die gegen x_0 konvergierte; es liesse sich dann wegen der Kompaktheit der Menge $\tau^{-1}(x_0) \cup \bigcup_{\nu=1}^{\infty} \tau^{-1}(x_\nu)$ in $\bigcup_{\nu=1}^{\infty} \tau^{-1}(x_\nu)$ eine Punktfolge $'x_\mu$ wählen, die gegen einen Punkt von $\tau^{-1}(x_0)$ konvergierte; da aber $\tau^{-1}(x_0)$ nur aus $'x_0$ besteht, würden fast alle der $'x_\mu$ in $'V$, also unendlich viele der x_ν in A liegen, im Widerspruch zur Wahl der x_ν. Es folgt, dass es eine in A enthaltene offene Umgebung $V_0 = V_0(x_0)$ gibt, sodass $\tau^{-1} \mid V_0 \colon V_0 \xrightarrow[k]{} {'X}$ eine holomorphe Abbildung ist; demnach ist x_0 eine gewöhnlicher Punkt inbezug auf die eigentliche Modifikation $('X, \tau, X)$. Die Menge der singulären Punkte inbezug auf $('X, \tau, X)$ stimmt also mit der analytischen Menge N überein. Damit ist Proposition 4.1.1 bewiesen, und es ergibt sich auch die Richtigkeit von Proposition 4.1.2.

Das Bild eines Punktes $x \in X$ unter einer meromorphen Abbildung

$f\colon X \xrightarrow[m]{} Y$ ist stets eine nichtleere kompakte analytische Menge in Y. Unter besonderen Voraussetzungen lassen sich über die Bildmengen $f(x)$ genauere Aussagen machen.

PROPOSITION 4.1.3. *Sei $f\colon X \xrightarrow[m]{} Y$ eine meromorphe Abbildung. Ist X im Punkte x irreduzibel, so ist $f(x)$ zusammenhängend. Ist X eine normaler komplexer Raum, so ist $x \in X$ genau dann eine ausserwesentliche Singularität von f, wenn $\dim f(x) > 0$ ist.*

Beweis. $f(x) \subset Y$ ist genau dann zusammenhängend, wenn $\pi[f]^{-1}(x) \subset G[f]$ zusammenhängend ist. Es sei L eine zusammenhängende Komponente von $\pi[f]^{-1}(x)$. Dann lassen sich je eine offene zusammenhängende Umgebung V von x und V^* von L so wählen, dass $\pi[f](V^*) = V$ gilt, dass die durch Beschränkung von $\pi[f]$ bestimmte Abbildung $\pi^*\colon V^* \to V$ eigentlich ist (vgl. [29], Hilfssatz 3)[8] und dass $V^* \cap \pi[f]^{-1}(x) = L$ ist. Ist nun X in x irreduzibel, so kann V ausserdem als irreduzibel angenommen werden (nötigenfalls sind V and V^* noch geeignet zu verkleinern). Es ist $r(\pi^*) = \dim V^* - \dim X$. $\pi^*(V^*)$ ist demnach eine analytische Menge von maximaler Dimension in V, wegen der Irreduzibilität von V ist also $\pi^*(V^*) = V$. Bezeichnet U die Menge der inbezug auf f ausgezeichneten Punkte in V, so folgt, dass $\pi[f]^{-1}(U) = \pi^{*-1}(U) \subset V^*$ gilt. Andererseits liegt $\pi[f]^{-1}(U)$ sicher in der offenen Menge $\pi[f]^{-1}(V)$ dicht, mithin ist wegen $V^* \subset \pi[f]^{-1}(V)$ notwendig $V^* = \pi[f]^{-1}(V)$. Daher ist auch $L = \pi[f]^{-1}(x)$, d.h. $\pi[f]^{-1}(x)$ ist zusammenhängend.

Ist X normal (also insbesondere lokal-irreduzibel), so sei $x_0 \in X$ ein Punkt, derart, dass $\dim f(x_0) = 0$ ist. $f(x_0)$ ist dann ein Punkt—denn $f(x_0)$ ist, wie soeben bewiesen, zusammenhängend—, also ist auch $\pi[f]^{-1}(x_0)$ ein Punkt $x_0^* \in G[f]$. Der lokale Rang von $\pi[f]$ in x_0 ist maximal (d.h. gleich $\dim G[f] = \dim X$); gleiches gilt dann für alle Punkte einer offenen Umgebung U_0^* von x_0^*. Es folgt, dass $U_0 := \pi[f](U_0^*)$ ebenfalls offen, also eine Umgebung von x_0 ist. Bezeichnet $\pi_0\colon U_0^* \to U_0$ die durch Beschränkung von $\pi[f]$ bestimmte Abbildung, so ist für $x' \in U_0$ stets $\pi_0^{-1}(x')$ eine nichtleere diskrete Menge, und diese muss einpunktig sein, da sonst $\pi[f]^{-1}(x')$ nicht zusammenhängend wäre und dann gleiches für $f(x')$ gälte. Demnach ist π_0

---

[8] Der in [29], p. 77, gegebene Beweis des Hilfssatzes 3 hat eine Lücke: Es ist nicht ohne weiteres ersichtlich, dass es eine Umgebung V der Zusammenhangskomponente N von $F^{-1}(F(p))$ mit der dort beschriebenen Eigenschaft gibt. Dies lässt sich jedoch zeigen. Wird noch vorausgesetzt, dass $F^{-1}(F(p)) - N$ in $F^{-1}(F(p))$ abgeschlossen ist, so ist die Existenz eines V evident. Diese zusätzliche Voraussetzung ist bei allen Anwendungen des Hilfssatzes 3 in [29] erfüllt und auch im obigen Falle gegeben.

bijektiv, also wegen der Normalität von X sogar biholomorph. Für $x' \in U_0$ gilt $\pi[f]^{-1}(x') = \pi_0^{-1}(x')$, daher ist $f(x') = f \circ \pi_0^{-1}(x')$, woraus sich ergibt, dass $f \mid U_0 : U_0 \to Y$ eine holomorphe Abbildung und infolgedessen x_0 ein nichtsingulärer Punkt inbezug auf f ist. Für jede ausserwesentliche Singularität $x \in X$ muss mithin $\dim f(x) > 0$ sein. Dass umgekehrt jeder Punkt $x \in X$, für welchen $\dim f(x) > 0$ gilt, eine ausserwesentliche Singularität von f ist, ist klar.

4.2. Wir beschränken die Betrachtung meromorpher Abbildungen von nun an der Einfachheit halber auf den Fall irreduzibler komplexer Urbildräume. Die im folgenden in diesem und den weiteren Paragraphen erörterten und mit meromorphen Abbildungen zusammenhängenden Begriffsbildungen und Aussagen nebst Beweisen lassen sich jedoch ohne Schwierigkeit mutatis mutandis auf den Fall reduzibler Urbildräume übertragen.

PROPOSITION 4.2.1. *Es seien* $f: X \underset{m}{\longrightarrow} X$ *und* $f_1 : X_1 \underset{m}{\longrightarrow} Y_1$ *meromorphe Abbildungen. Dann ist auch* $f \times f_1 : X \times X_1 \underset{m}{\longrightarrow} Y \times Y_1$ *eine meromorphe Abbildung.*

Beweis. Nach Proposition 3.4.1 ist $f \times f_1$ eine holomorphe Korrespondenz. Die Irreduzibilität des Graphen $G[f \times f_1]$ ergibt sich daraus, dass mit $G[f]$ und $G[f_1]$ auch $G[f] \times G[f_1]$ ein irreduzibler komplexer Raum ist. Ist ferner $x \in X$ ein ausgezeichneter Punkt inbezug auf f, $x_1 \in X_1$ ein ausgezeichneter Punkt inbezug auf f_1, so ist $(x, x_1) \in X \times X_1$ ein ausgezeichneter Punkt inbezug auf $f \times f_1$.

Die Verbindung zweier meromorpher Abbildungen $f: X \underset{m}{\longrightarrow} Y$ und $f_1' : X \underset{m}{\longrightarrow} Y_1$ ist im allgemeinen nicht wieder eine meromorphe Abbildung, da—wie einfache Beispiele zeigen—der Graph von $(f, f_1') : X \underset{k}{\longrightarrow} Y \times Y_1$ nicht notwendig irreduzibel sein muss. Ist weiter $g: Y \underset{m}{\longrightarrow} Z$ eine meromorphe Abbildung, so braucht auch das Produkt $g \circ f : X \underset{k}{\longrightarrow} Z$ keine meromorphe Abbildung zu sein. Es lassen sich aber eine Verbindung zweiter Art und unter einer besonderen Voraussetzung ein Produkt zweiter Art für meromorphe Abbildungen so einführen, dass diese Verknüpfungen wieder meromorphe Abbildungen ergeben.—Wir benötigen

PROPOSITION 4.2.2. *Sei* $f' : X \underset{k}{\longrightarrow} Y$ *eine holomorphe Korrespondenz des irreduziblen komplexen Raumes* X *in den komplexen Raum* Y. *Es existiere in* X *ein ausgezeichneter Punkt inbezug auf* f'. *Dann gibt es genau* *eine meromorphe Abbildung $f\colon X \xrightarrow[m]{} Y$, die in f' enthalten ist. Es gibt ferner eine von X verschiedene analytische Menge M in X, derart, dass $f' \mid X - M$ eine holomorphe Abbildung und daher $f' \mid X - M = f \mid X - M$ ist.*

Beweis. Ist $x_0 \in X$ ein ausgezeichneter Punkt inbezug auf f', so sei f diejenige irreduzible Komponente von f', deren Graph $G[f]$ den Punkt $\pi[f']^{-1}(x_0)$ enthält. x_0 ist auch ausgezeichneter Punkt inbezug auf f, daher ist f eine meromorphe Abbildung. Ist f_0 eine weitere in f' enthaltene meromorphe Abbildung, so ist jeder inbsezug auf f' ausgezeichnete Punkt auch ausgezeichneter Punkt inbezug auf f_0. In einer Umgebung von x_0 stimmt daher f' mit f_0 und mit f überein; nach Proposition 3.4.2 ist dann $f_0 = f$. Hat f' ausser f noch weitere irreduzible Komponenten $f^{(j)}$, $j \in J$, so sei $f'' := \bigcup_j f^{(j)}$; es ist dann $\pi[f''](G[f'']) =: N$ eine analytische Menge in X, und diese enthält x_0 nicht, ist also von X verschieden. Ist andererseits $f' = f$, so sei $N := \emptyset$. Es hat dann $M := N \cup S(f)$ die oben angegebene Eigenschaft.

Sind nun $f_1\colon X \xrightarrow[m]{} Y_1$, $f_2\colon X \xrightarrow[m]{} Y_2$ meromorphe Abbildungen, so ist das in X dichte Gebiet $U := X - (S(f_1) \cup S(f_2))$ ausgezeichnet inbezug auf die holomorphe Korrespondenz $(f_1, f_2)\colon X \xrightarrow[k]{} Y_1 \times Y_2$. Die nach Proposition 4.2.2 in (f_1, f_2) enthaltene eindeutig bestimmte meromorphe Abbildung sei als *Verbindung zweiter Art* oder *meromorphe Verbindung* von f_1 und f_2 bezeichnet; wir kennzeichnen sie durch die Schreibweise $\{f_1, f_2\}\colon X \xrightarrow[m]{} Y_1 \times Y_2$.

Für drei meromorphe Abbildungen $f_i\colon X \xrightarrow[m]{} Y_i$ $(i = 1, 2, 3)$ sind die meromorphen Verbindungen $\{\{f_1, f_2\}, f_3\}$ und $\{f_1, \{f_2, f_3\}\}$ definiert und, wie sich mit Benutzung von Proposition 3.4.2 ergibt, gleich; es darf also $\{f_1, f_2, f_3\} := \{\{f_1, f_2\}, f_3\} = \{f_1, \{f_2, f_3\}\}$ geschrieben werden. Allgemein ist dann auch für l meromorphe Abbildung $f_\lambda\colon X \xrightarrow[m]{} Y_\lambda$ $(\lambda = 1, \cdots, l; l \geqq 2)$ die meromorphe Verbindung $\{f_1, \cdots, f_l\}\colon X \xrightarrow[m]{} Y_1 \times \cdots \times Y_l$ definiert.

Es seien weiter meromorphe Abbildungen $f\colon X \xrightarrow[m]{} Y$ und $g\colon Y \xrightarrow[m]{} Z$ gegeben. Falls es in X einen inbezug auf die holomorphe Korrespondenz $g \circ f\colon X \xrightarrow[k]{} Z$ ausgezeichneten Punkt gibt, so existiert nach Proposition 4.2.2 genau eine in $g \circ f$ enthaltene meromorphe Abbildung; wir nennen dann diese meromorphe Abbildung das *Produkt zweiter Art* oder das *meromorphe*

Produkt von f und g und schreiben $g \triangle f\colon X \xrightarrow[m]{} Z$.[9] Wenn im folgenden die Sprechweise benutzt wird, dass $g \triangle f$ definiert sei, so ist stets gemeint, dass $g \circ f$ die obige Voraussetzung erfüllt.

Das meromorphe Produkt $g \triangle f$ ist insbesondere dann definiert, wenn die Bildmenge $f(X)$ nicht in der Menge $S(g)$ der ausserwesentlichen Singularitäten von g enthalten ist. Denn trifft dies zu, so ist die Menge $U\colon = X - (S(f) \cup f^{-1}(S(g)))$ nichtleer, und für jeden Punkt $x \in U$ ist die Menge $g \circ f(x)$ einpunktig; also ist U eine inbezug auf $g \circ f$ ausgezeichnete Menge in X. Ist f surjektiv, so ist $f(X)$ nicht in $S(g)$ enthalten, also ist dann $g \triangle f$ immer definiert.

In besonderen Fällen stimmen $g \circ f$ und $g \triangle f$ überein. Z. B. gilt: *Ist $f\colon X \xrightarrow[m]{} Y$ eine meromorphe und $g\colon Y \longrightarrow Z$ eine holomorphe Abbildung, so ist $g \circ f = g \triangle f$.* Denn der Graph $G[g \circ f]$ lässt sich aus dem Graphen $G[g \circ \hat{f}] \subset G[f] \times Z$ der holomorphen Abbildung $g \circ \hat{f}\colon G[f] \to Z$ vermöge der eigentlichen holomorphen Abbildung

$$\pi[f] \times \iota_Z\colon G[f] \times Z \to X \times Z \quad (\iota_Z\colon Z \to Z \text{ die identische Abbildung})$$

erhalten, mit $G[g \circ \hat{f}]$ ist dann auch $G[g \circ f] = (\pi[f] \times \iota_Z)(G[g \circ \hat{f}])$ irreduzibel.—Ist $f\colon X \to Y$ holomorph und $g\colon Y \xrightarrow[m]{} Z$ meromorph, so gilt, falls $g \triangle f$ definiert ist, im allgemeinen $g \circ f \neq g \triangle f$. Z.B. ist, wenn g Unbestimmtheitsstellen besitzt, stets $g \triangle \pi[g] = \hat{g}$ in $g \circ \pi[g]$ echt enthalten.

Ist A eine nichtleere irreduzible lokal-analytische Menge in X, $\epsilon\colon A \to X$ die Injektion und $f\colon X \xrightarrow[m]{} Y$ so beschaffen, dass $f \triangle \epsilon\colon A \xrightarrow[m]{} Y$ definiert ist, so nennen wir $f \triangle \epsilon$ die *meromorphe Beschränkung* von f auf A und schreiben $f \underset{m}{|} A\colon = f \triangle \epsilon$. Es ist stets $r(f \underset{m}{|} A) \leqq r(f)$, wie sich mittels Proposition 1.2.1 ergibt.

Im allgemeinen ist $f \underset{m}{|} A$ von der holomorphen Korrespondenz $f \mid A = f \circ \epsilon\colon A \xrightarrow[k]{} Y$ verschieden, da der Graph $G[f \mid A]$ nicht irreduzibel zu sein braucht. Ein Fall, in dem aber $f \underset{m}{|} A$ mit $f \mid A$ zusammenfällt, liegt vor, wenn A ein Gebiet in X ist: Es ist $A_0\colon = A - S(f)$ ein in A dichtes irreduzibles Gebiet; dann ist auch $\pi[f]^{-1}(A_0)$ ein in $\pi[f]^{-1}(A)$ dichtes irreduzibles Gebiet, woraus sich die Irreduzibilität von $G[f \mid A]$ und damit die Übereinstimmung von $f \mid A$ und $f \underset{m}{|} A$ ergibt.

Dass für eine lokal-analytische Menge A' in Y, in der die Bildmenge

---

[9] Vgl. hierzu und zu folgenden Ausführungen auch [16] und [33].

$f(X)$ enthalten ist, die Beschränkung von f auf A' stets wieder eine meromorphe Abbildung ist, ist klar.

Für drei meromorphe Abbildungen $f: X \underset{m}{\longrightarrow} Y, g: Y \underset{m}{\longrightarrow} Z, h: Z \underset{m}{\longrightarrow} Z'$ kann $h \triangle (g \triangle f)$ definiert sein, ohne dass $(h \triangle g) \triangle f$ es ist, und umgekehrt. Die meromorphen Abbildungen sind aber dann beide definiert und gleich, wenn f und g surjektiv sind. Denn in diesem Falle sind zunächst $g \triangle f$ und $(h \triangle g) \triangle f$ definiert; es ist ferner

$$U: = X - (S(f) \cup f^{-1}(S(g)) \cup (g \circ f)^{-1}(S(h)))$$

eine inbezug auf $h \circ g \circ f: X \underset{k}{\longrightarrow} Z'$ ausgezeichnete offene Menge in X. Für $x \in U$ gilt dann

$$h \circ (g \triangle f)(x) \subset h \circ g \circ f(x) \text{ und } (h \triangle g) \triangle f(x) \subset h \circ g \circ f(x),$$

also—da die Menge $h \circ g \circ f(x)$ einpunktig ist— $h \circ (g \triangle f)(x) = h \circ g \circ f(x) = (h \triangle g) \triangle f(x)$. Demnach ist U auch ausgezeichnet inbezug auf $h \circ (g \triangle f)$, mithin ist $h \triangle (g \triangle f)$ definiert; weiter stimmen $h \triangle (g \triangle f)$ und $(h \triangle g) \triangle f$ auf U überein, daher ist nach Proposition 3.4.2

$$h \triangle (g \triangle f) = (h \triangle g) \triangle f.$$

Es lässt sich leicht allgemein zeigen, dass, wenn die meromorphen Produkte $h \triangle (g \triangle f)$ und $(h \triangle g) \triangle f$ beide definiert sind, stets $h \triangle (g \triangle f) = (h \triangle g) \triangle f$ ist.

Eine meromorphe Abbildung $f: X \underset{m}{\longrightarrow} Y$ heisst *bimeromorph*, wenn $f^{-1}: Y \underset{m}{\longrightarrow} X$ eine meromorphe Abbildung ist.[10] Sei f bimeromorph. Dann ist f surjektiv und, da f^{-1} insbesondere stetig ist, eigentlich. Weiter ist f^{-1} mit $f = (f^{-1})^{-1}$ bimeromorph. Wird $G[f] \subset X \times Y$ vermöge der durch $(x, y) \rightarrow (y, x)$ $(x \in X,\ y \in Y)$ gegebenen biholomorphen Abbildung $X \times Y \rightarrow Y \times X$ mit $G[f^{-1}] \subset Y \times X$ identifiziert, was im folgenden geschehen soll, so hat man $\pi[f^{-1}] = \hat{f}$; also ist $f = \pi[f^{-1}] \circ \pi[f]^{-1}$ und $f^{-1} = \pi[f] \circ \pi[f^{-1}]^{-1}$. Das in X dichte Gebiet $U_X: = X - (S(f) \cup f^{-1}(S(f^{-1})))$ wird vermöge f biholomorph auf das in Y dichte Gebiet $U_Y: = Y - (S(f^{-1}) \cup f(S(f)))$ abgebildet und umgekehrt U_Y vermöge f^{-1} biholomorph auf U_X; denn $X - S(f)$ wird vermöge $\pi[f]^{-1}$ biholomorph auf $G[f] - \pi[f]^{-1}(S(f))$ bezogen und $Y - S(f^{-1})$ vermöge $\pi[f^{-1}]^{-1}$ biholomorph auf auf $G[f] - \pi[f^{-1}]^{-1}(S(f^{-1}))$.

---

[10] Diese Erklärung gilt auch für reduzible Räume X, Y. Ist X irreduzibel, was oben gemäss der Verabredung auf p. 838 vorauszusetzen ist, so ist Y, wenn eine bimeromorphe Abbildung $f: X \underset{m}{\longrightarrow} Y$ vorliegt, von selbst ebenfalls irreduzibel.

Demnach ist U_X ausgezeichnet inbezug auf $f^{-1} \circ f$, und für $x \in U_X$ ist $f^{-1} \circ f(x) = x$; also ist $f^{-1} \triangle f \colon X \xrightarrow[m]{} X$ nach Proposition 3.4.2 die identische Abbildung.

Sei $g \colon Y \xrightarrow[m]{} Z$ eine weitere bimeromorphe Abbildung, dann sind $g \triangle f \colon X \xrightarrow[m]{} Z$ und $f^{-1} \triangle g^{-1} \colon Z \xrightarrow[m]{} X$ definiert. Die Korrespondenz $(g \triangle f)^{-1} \colon Z \xrightarrow[k]{} X$ ist jedenfalls holomorph und irreduzibel, denn der Graph von $(g \triangle f)^{-1}$ ist als irreduzible analytische Untermenge im Graphen der holomorphen Korrespondenz $(g \circ f)^{-1} = f^{-1} \circ g^{-1} \colon Z \xrightarrow[k]{} X$ enthalten. Es wird nun $U_X' := X - (S(f) \cup f^{-1}(S(f^{-1})) \cup f^{-1}(S(g)) \cup f^{-1} \circ g^{-1}(S(g^{-1})))$ vermöge $g \circ f \colon X \xrightarrow[k]{} Z$ biholomorph auf

$$U_Z' = Z - (S(g^{-1}) \cup g(S(g)) \cup g(S(f^{-1})) \cup g \circ f(S(f)))$$

abgebildet und umgekehrt U_Z' vermöge $(g \circ f)^{-1}$ biholomorph auf U_X' (U_X' ist ein in X dichtes Gebiet und U_Z' ein in Z dichtes Gebiet). Daher ist U_X' ausgezeichnet inbezug auf $g \circ f$ und U_Z' ausgezeichnet inbezug auf $(g \circ f)^{-1}$. Sei z beliebig in U_Z' gewählt und $(g \circ f)^{-1}(z) =: x$. Wegen $g \triangle f(x) = g \circ f(x) = z$ ist dann auch $(g \triangle f)^{-1}(z) = x$, und es ist ebenfalls $f^{-1} \triangle g^{-1}(z) = f^{-1} \circ g^{-1}(z) = x$. Demnach ist U_Z' ausgezeichnet inbezug auf $(g \triangle f)^{-1}$, und es folgt, dass $(g \triangle f)^{-1}$ eine meromorphe Abbildung, dass also $g \triangle f$ mit g und f bimeromorph ist. Wegen $(g \triangle f)^{-1} \mid U_Z' = f^{-1} \triangle g^{-1} \mid U_Z'$ ist nach Proposition 3.4.2 ferner $(g \triangle f)^{-1} = f^{-1} \triangle g^{-1}$.

Wir stellen noch fest, dass jede eigentliche Modifikationsabbildung bimeromorph ist.

4.3. PROPOSITION 4.3.1. *Es seien* $f \colon X \xrightarrow[m]{} Y, f_1 \colon X \xrightarrow[m]{} Y_1$ *meromorphe Abbildungen, f sei eigentlich. Dann sind auch* $\{f, f_1\} \colon X \xrightarrow[m]{} Y \times Y_1$ *und* $\{f_1, f\} \colon X \xrightarrow[m]{} Y_1 \times Y$ *eigentliche meromorphe Abbildungen.*

Beweis. Nach Proposition 3.3.2 sind die holomorphen Korrespondenzen $(f, f_1) \colon X \xrightarrow[k]{} Y \times Y_1$ und $(f_1, f) \colon X \xrightarrow[k]{} Y_1 \times Y$ eigentlich. Dann sind aber auch alle in (f, f_1) bezw. (f_1, f) enthaltenen holomorphen Korrespondenzen eigentlich, insbesondere $\{f, f_1\}$ und $\{f_1, f\}$.

PROPOSITION 4.3.2. *Es seien* $f \colon X \xrightarrow[m]{} Y$ *und* $g \colon Y \xrightarrow[m]{} Z$ *meromorphe Abbildungen, derart, dass* $g \triangle f \colon X \xrightarrow[m]{} Z$ *definiert ist. Dann gilt:*

(a) *Sind f und g eigentlich, so ist auch $g \triangle f$ eigentlich.*

(b) *Ist $g \triangle f$ eigentlich, so ist f eigentlich. Ist überdies f surjektiv, so ist auch g eigentlich.*

Beweis. (a) Die holomorphe Korrespondenz $g \circ f \colon X \xrightarrow[k]{} Z$ ist eigentlich (Proposition 3.3.1); gleiches trifft dann für die in $g \circ f$ enthaltene meromorphe Abbildung $g \triangle f$ zu.

(b) Es sei $U \subset X$ eine inbezug auf $g \circ f \colon X \xrightarrow[k]{} Z$ ausgezeichnete offene Menge. Für jeden Punkt $u \in U$ ist mit $g \circ f(u)$ stets $\pi[g]^{-1} \circ f(u)$ eine einpunktige Menge; U ist daher auch ausgezeichnet inbezug auf $\pi[g]^{-1} \circ f =: X \xrightarrow[k]{} G[g]$. Mithin ist das meromorphe Produkt $\pi[g]^{-1} \triangle f =: \phi \colon X \xrightarrow[m]{} G[g]$ definiert. Für $u \in U$ gilt nun

$$g \triangle f(u) = g \circ f(u) = \hat{g} \circ \pi[g]^{-1} \circ f(u) = \hat{g} \circ \phi(u) = \hat{g} \circ \hat{\phi} \circ \pi[\phi]^{-1}(u);$$

auf Grund von Proposition 3.4.2 ist also

$$g \triangle f = (\hat{g} \circ \hat{\phi}) \triangle \pi[\phi]^{-1} = \hat{g} \circ \hat{\phi} \circ \pi[\phi]^{-1}.$$

Es ist ferner

$$\pi[g] \circ \phi(u) = \pi[g] \circ \pi[g]^{-1} \circ f(u) = f(u),$$

mithin $\pi[g] \circ \hat{\phi} \circ \pi[\phi]^{-1}(u) = f(u)$, woraus sich $\pi[g] \circ \hat{\phi} \circ \pi[\phi]^{-1} = f$ ergibt.

Da $g \triangle f$ als eigentlich vorausgesetzt ist, folgt wegen $g \triangle f = \hat{g} \circ \hat{\phi} \circ \pi[\phi]^{-1}$ aus Proposition 3.3.2, dass $\hat{g} \circ \hat{\phi}$ und somit $\hat{\phi}$ eigentlich ist. Wegen $\pi[g] \circ \hat{\phi} \circ \pi[\phi]^{-1} = f$ ist dann nach Proposition 3.3.1 auch f eigentlich.

Ist f surjektiv, so muss auch $\hat{\phi}$ surjektiv sein. Denn träfe dies nicht zu, so wäre $\hat{\phi}(G[\phi])$ eine von $G[g]$ verschiedene analytische Menge in $G[g]$, und es wäre dann $\pi[g] \circ \hat{\phi}(G[\phi]) = \pi[g] \circ \hat{\phi} \circ \pi[\phi]^{-1}(X) = f(X)$ von Y verschieden. Aus der Eigentlichkeit von $\hat{g} \circ \hat{\phi}$ und der Surjektivität von $\hat{\phi}$ folgt nun, dass $\hat{g}$ und mithin g eigentlich ist.

5. Maximale meromorphe Abbildungen.

5.1. Sei $f \colon X \xrightarrow[m]{} Y$ eine meromorphe Abbildung.[11] Der Rang $r(f)$ von f ist gemäss 3.4 definiert als der Rang $r(\hat{f})$ der f durch Liftung nach $G[f]$ zugeordneten holomorphen Abbildung $\hat{f} \colon G[f] \to Y$. Ist f insbesondere eine holomorphe Abbildung, so ist $r(f)$ der globale Rang von f in dem für holomorphe Abbildungen definierten Sinne. Für jeden offenen irreduziblen komplexen Unterraum X_0 von X gilt $r(f \mid X_0) = r(f)$.

---

[11] Es sei daran erinnert, dass—wie in 4.2 festgelegt—Urbildräume meromorpher Abbildungen stets als irreduzibel vorausgesetzt werden.

Für zwei meromorphe Abbildung $f\colon X \xrightarrow[m]{} Y$ und $f_0\colon X \xrightarrow[m]{} Y_0$ ist stets $r(\{f,f_0\}) \geqq \max(r(f), r(f_0))$. Wir sagen, f_0 sei von f *abhängig*, wenn $r(\{f,f_0\}) = r(f)$ gilt (insbesondere ist dann $r(f) \geqq r(f_0)$). Ist f_0 von f abhängig und f von f_0 abhängig, so heissen f und f_0 *analytisch verwandt*.—Es lässt sich leicht nachweisen, dass folgendes gilt:

(a) *Sind die Abbildungen* $f_0\colon X \xrightarrow[m]{} Y_0$ *und* $f_1\colon X \xrightarrow[m]{} Y_1$ *von* f *abhängig, so ist auch* $\{f_0, f_1\}\colon X \xrightarrow[m]{} Y_0 \times Y_1$ *von* f *abhängig, und* $\{f, f_0\}$ *sowie* $\{f, f_1\}$ *sind mit* f *analytisch verwandt.*

(b) *Ist* $f_0\colon X \xrightarrow[m]{} Y_0$ *von* f *abhängig und* $'f_0\colon X \xrightarrow[m]{} {'Y_0}$ *von* f_0 *abhängig, so ist auch* $'f_0$ *von* f *abhängig.*

Es gilt ferner

(c) *Ist* $g\colon Y \xrightarrow[m]{} Z$ *eine meromorphe Abbildung, derart, dass* $g \triangle f\colon X \xrightarrow[m]{} Z$ *definiert ist, so ist* $g \triangle f$ *von* f *abhängig.*

Beweis. Wegen Proposition 4.2.2 lässt sich in X ein dort dichtes irreduzibles Gebiet V wählen, derart, dass $f \mid V$ und $g \circ f \mid V$ holomorphe Abbildungen sind. Da $g(f(v))$ für jeden Punkt $v \in V$ eine einpunktige Menge ist, bleibt $g \circ f \mid V$ auf den Fasern von $f \mid V$ konstant. Es folgt, dass $g \circ f \mid V$ von $f \mid V$ abhängig, also $r(f \mid V, g \circ f \mid V) = r(f \mid V)$ ist. Wegen

$$(f \mid V, g \circ f \mid V) = (f, g \circ f) \mid V = \{f, g \triangle f\} \mid V,$$
$$r(\{f, g \triangle f\} \mid V) = r(\{f, g \triangle f\}) \text{ und } r(f \mid V) = r(f)$$

ist dann auch $r(\{f, g \triangle f\}) = r(f)$.

Ist $f\colon X \xrightarrow[m]{} Y$ eigentlich, wird ausser X auch Y als irreduzibel vorausgesetzt, und ist $\dim X = \dim Y = r(f)$, so lässt sich ein Abbildungsgrad von f erklären: Sei E die Entartungsmenge von $\hat{f}\colon G[f] \to Y$, es werde $M_1'\colon = \pi[f](E) \cup S(f)$, $M_2 \colon = f(M_1')$, $M_1 \colon = f^{-1}(f(M_1'))$ gesetzt. M_1, M_2 sind von X bezw. Y verschiedene analytische Mengen, es gilt $f(M_1) = M_2$, $f^{-1}(M_2) = M_1$, und die durch Beschränkung von f bestimmte holomorphe Abbildung $f'\colon X - M_1 \to Y - M_2$ ist eine eigentliche Überlagerungsabbildung. Die Blätterzahl der analytischen Überlagerung $(X - M_1, f', Y - M_2)$ sei nun als *Abbildungsgrad* $\mathfrak{g}(f)$ von f bezeichnet. Wir nennen das Tripel (X, f, Y) eine *meromorphe Überlagerung* oder *m-Überlagerung* von Y und f eine *eigentliche meromorphe Überlagerungsabbildung*; $\mathfrak{g}(f)$ heisst die *Blätterzahl* der Überlagerung.

Es gelten die leicht nachzuweisenden Aussagen:

(I) *Sind* $f\colon X \xrightarrow[m]{} Y$ *und* $g\colon Y \xrightarrow[m]{} Z$ *eigentliche meromorphe Überlagerungsabbildungen, so ist* $g \triangle f\colon X \xrightarrow[m]{} Z$ *definiert und eine eigentliche meromorphe Überlagerungsabbildung mit* $\mathfrak{g}(g \triangle f) = \mathfrak{g}(g) \cdot \mathfrak{g}(f)$.

(II) *Eine meromorphe Abbildung* $f\colon X \xrightarrow[m]{} Y$ *ist genau dann bimeromorph, wenn* f *eine eigentliche meromorphe Überlagerungsabbildung mit* $\mathfrak{g}(f) = 1$ *ist.*

Als Folgerung hieraus ergibt sich: *Ist* $f\colon X \xrightarrow[m]{} Y$ *surjektiv und* $g\colon Y \xrightarrow[m]{} X$ *so beschaffen, dass* $g \triangle f\colon X \xrightarrow[m]{} X$ *die identische Abbildung von* X *ist, so sind* f *und* g *bimeromorph, und es ist* $g = f^{-1}$. Denn nach Proposition 4.3.2 sind f und g eigentlich; es ist ferner $\dim X \geqq r(f) = \dim Y$ und $\dim Y \geqq r(g) = \dim X$, also $\dim X = \dim Y = r(f) = r(g)$. Demnach sind f und g eigentliche meromorphe Überlagerungsabbildungen; wegen $\mathfrak{g}(g \triangle f) = 1$ gilt $\mathfrak{g}(f) = \mathfrak{g}(g) = 1$. Also sind f und g bimeromorph, und es ist $g = g \triangle (f \triangle f^{-1}) = (g \triangle f) \triangle f^{-1} = f^{-1}$.

5.2. In Analogie zu Proposition 1.2.2 gilt

PROPOSITION 5.2.1. *Sei* $f\colon X \xrightarrow[m]{} Y$ *eine meromorphe Abbildung,* A *eine irreduzible analytische Menge in* X*, derart, dass* $f \underset{m}{\mid} A$ *definiert und eigentlich und dass* $r(f) = r(f \underset{m}{\mid} A)$ *ist. Dann ist* $(f \underset{m}{\mid} A)(A) = f(A) = f(X)$, *und* $f(X)$ *ist eine* $r(f)$*-dimensionale irreduzible analytische Menge in* Y.

Beweis. Es sei $A_0 \subset A$ ein irreduzibles und inbezug auf $f \mid A\colon A \xrightarrow[k]{} Y$ ausgezeichnetes Gebiet. Vermöge $\pi[f]^{-1}$ wird A_0 auf ein irreduzibles Gebiet $'A_0$ in der analytischen Menge $\pi[f]^{-1}(A) \subset G[f]$ bijektiv bezogen. Sei $'A$ diejenige irreduzible Komponente von $\pi[f]^{-1}(A)$, die $'A_0$ umfasst, weiter $'\alpha\colon {'A} \to A$ die durch Beschränkung von $\pi[f]$ bestimmte holomorphe Abbildung. $'\alpha$ ist eigentlich, und $'A_0$ wird auf A_0 vermöge $'\alpha$ bijektiv abgebildet; infolgedessen ist $'\alpha$ eine eigentliche Modifikationsabbildung. Die meromorphe Abbildung $(\hat{f} \mid {'A}) \triangle {'\alpha}^{-1} = (\hat{f} \mid {'A}) \circ {'\alpha}^{-1}\colon A \xrightarrow[m]{} Y$ stimmt auf A_0 mit $f \underset{m}{\mid} A$ überein, daher ist $(\hat{f} \mid {'A}) \circ {'\alpha}^{-1} = f \underset{m}{\mid} A$. Da $f \underset{m}{\mid} A$ eigentlich ist, hat $\hat{f} \mid {'A}$

nach Proposition 3.3.2 diese Eigenschaft auch; ferner ist $r(f \underset{m}{|} A) = r(\hat{f} \mid {'A})$ und $r(f) = r(\hat{f})$, also $r(\hat{f} \mid {'A}) = r(\hat{f})$. Es ist nun

$$(f \underset{m}{|} A)(A) = (\hat{f} \mid {'A})({'\alpha^{-1}}(A)) = (\hat{f} \mid {'A})({'A}) = \hat{f}({'A}),$$

und $f(X) = \hat{f}(\pi[f]^{-1}(X)) = \hat{f}(G[f])$. Nach Proposition 1.2.2 ist $\hat{f}({'A}) = \hat{f}(G[f])$, mithin ist $(f \underset{m}{|} A)(A) = f(X)$ und $f(X) = \hat{f}({'A})$ eine irreduzible analytische Menge der Dimension $r(\hat{f}) = r(f)$. Wegen $(f \underset{m}{|} A)(A) \subset f(A) \subset f(X)$ ist dann auch $(f \underset{m}{|} A)(A) = f(A) = f(X)$.

Wir benötigen weiter

PROPOSITION 5.2.2. *Sei* $f\colon X \xrightarrow[m]{} Y$ *eine meromorphe Abbildung,* A *eine irreduzible lokal-analytische Menge in* X*, derart, dass* X *in allen Punkten von* A *irreduzibel und dass* $f \underset{m}{|} A$ *definiert und* $r(f) = r(f \underset{m}{|} A)$ *ist. Dann ist für jede von* f *abhängige meromorphe Abbildung* $f_0\colon X \xrightarrow[m]{} Y_0$ *die meromorphe Beschränkung* $f_0 \underset{m}{|} A\colon A \xrightarrow[m]{} Y_0$ *definiert.*

Beweis. 1) Wir setzen den komplexen Raum X zunächst als normal voraus. (X ist dann insbesondere in allen seinen Punkten irreduzibel.) In diesem Falle ist auf Grund von Proposition 4.1.3 die meromorphe Beschränkung auf A einer meromorphen Abbildung genau dann definiert, wenn es auf A wenigstens einen nichtsingulären Punkt inbezug auf die Abbildung gibt. Deswegen darf angenommen werden, dass f eine holomorphe Abbildung ist: Nötigenfalls ist X durch eine hinreichend kleine zusammenhängende offene Umgebung eines inbezug auf f nichtsingulären Punktes von A zu ersetzen.—Es ist zu zeigen, dass A nicht in der analytischen Menge $S(f_0) = : S_0$ der ausserwesentlichen Singularitäten von f_0 enthalten ist.

Angenommen, es gälte $A \subset S_0$. Dann sei S_0' eine irreduzible Komponente von S_0, die A enthält; wegen $r(f \mid A) \leqq r(f \mid S_0') \leqq r(f)$ und $r(f) = r(f \mid A)$ gilt $r(f) = r(f \mid S_0')$. Die Menge $\pi[f_0]^{-1}(S_0')$ ist analytisch in $G[f]$, und die durch Beschränkung von $\pi[f_0]$ bestimmte holomorphe Abbildung $\pi[f_0]^{-1}(S_0') \to S_0'$ ist eigentlich und surjektiv; es gibt dann auch eine irreduzible Komponente ${'S_0'}$ von $\pi[f_0]^{-1}(S_0')$, die vermöge $\pi[f_0]$ eigentlich *auf* S_0' abgebildet wird. Es ist $r(\pi[f_0] \mid {'S_0'}) < \dim {'S_0'}$. Denn für jeden Punkt ${'x_0} \in \pi[f_0]^{-1}(S_0')$ ist die Faser $\pi[f_0]^{-1}(\pi[f_0]({'x_0}))$ eine mindestens eindimensionale zusammenhängende analytische Menge, da entsprechendes für $f_0(\pi[f_0]({'x_0}))$ gilt und $\pi[f_0]^{-1}(\pi[f_0]({'x_0}))$ auf $f_0(\pi[f_0]({'x_0}))$ vermöge dere holomorphen Abbildung $\hat{f}_0\colon G[f_0] \to Y_0$ bijektiv bezogen wird. Ist $'x_0$ insbesondere ein Punkt von $'S_0'$, der auf keiner weiteren irreduziblen Komponente von $\pi[f_0]^{-1}(S_0')$ liegt, so muss die Vereingung der durch $'x_0$ laufenden irreduziblen Komponenten von $\pi[f_0]^{-1}(\pi[f_0]('x_0))$ in $'S_0'$ enthalten sein und daher mit der Faser von $\pi[f_0] \mid {}'S_0'$ durch $'x_0$ übereinstimmen; infolgedessen ist der lokale Rang von $\pi[f_0] \mid {}'S_0'$ in $'x_0$ kleiner als $\dim {}'S_0'$. Da die Menge derartiger Punkte $'x_0$ auf $'S_0'$ dicht liegt, gilt mithin $r(\pi[f_0] \mid {}'S_0)' < \dim {}'S_0'$. Jede Niveaumenge von $\pi[f_0] \mid {}'S_0'$ ist demnach mindestens ein dimensional.

Es werde weiter $'f := f \circ \pi[f_0]\colon G[f_0] \to Y$ gesetzt. Dann folgt, dass $r('f \mid {}'S_0') = r(f \mid S_0')$ ist. Denn ist B_0' eine offene nichtleere Teilmenge von S_0' mit abzählbarer Topologie, so ist $f(B_0')$ ein topologisch-$2r(f \mid S_0')$-dimensionaler topologischer Unterraum von Y (vgl. [26], Satz 1). Es ist weiter $\pi[f_0]^{-1}(B_0') \cap {}'S_0' =: {}'B_0'$ eine offene Menge in $'S_0'$ mit abzählbarer Topologie, und es gilt $\pi[f_0]('B_0') = B_0'$, also $'f('B_0') = f(B_0')$; da $'f('B_0')$ ein topologisch-$2r('f \mid {}'S_0')$-dimensionaler topologischer Unterraum von Y ist, muss also $r(f \mid S_0') = r('f \mid {}'S_0')$ sein.—Es ist $r(f) = r('f)$, demnach auch $r('f) = r('f \mid {}'S_0')$.

Da f_0 von f abhängig ist, ist $\hat{f}_0$ von f' abhängig. Die Abbildungen $'f$ und $('f, \hat{f}_0)\colon G[f_0] \to Y \times Y_0$ sind also analytisch verwandt. Gleiches muss auch für die Beschränkungen $'f \mid {}'S_0'$ und $('f, \hat{f}_0) \mid {}'S_0 = ('f \mid {}'S_0', \hat{f}_0 \mid {}'S_0)$ gelten. Denn es ist

$$r('f \mid {}'S_0') \leqq r(('f \mid {}'S_0', \hat{f}_0 \mid {}'S_0')) = r(('f, \hat{f}_0) \mid {}'S_0') \leqq r(('f, \hat{f}_0)),$$

wegen der analytischen Verwandtschaft von $'f$ und $('f, \hat{f}_0)$ ferner $r('f) = r(('f, \hat{f}_0))$ und wie oben festgestellt $r('f) = r('f \mid {}'S_0')$. Daher ist

$$r('f \mid {}'S_0') \leqq r(('f, \hat{f}_0) \mid {}'S_0') \leqq r(('f, \hat{f}_0)) = r('f) = r('f \mid {}'S_0'),$$

mithin $r('f \mid {}'S_0') = r(('f, \hat{f}_0) \mid {}'S_0')$. Nunmehr lässt sich wie folgt ein Widerspruch herbeiführen: $'f \mid {}'S_0'$ ist wegen $'f = f \circ \pi[f_0]$ auf den Fasern der holomorphen Abbildung $\pi[f_0] \mid {}'S_0'\colon {}'S_0' \to X$ konstant, also ist $'f \mid {}'S_0'$ von $\pi[f_0] \mid {}'S_0'$ abhängig. Dann müsste auch $('f, \hat{f}_0) \mid {}'S_0'$ von $\pi[f_0] \mid {}'S_0'$ abhängig sein, $('f, \hat{f}_0) \mid {}'S_0'$ müsste auf den Niveaumengen von $\pi[f_0] \mid {}'S_0'$ der kleinsten vorkommenden Dimension konstant bleiben. Dies kann jedoch nicht zutreffen: Denn $('f, \hat{f}_0) \mid {}'S_0'$ ist auf einer Niveaumenge von $\pi[f_0] \mid {}'S_0'$, die mehr als einen Punkt enthält, niemals konstant, da jede Faser von $\pi[f_0]$ vermöge $\hat{f}_0$ injektiv abgebildet wird; jede Niveaumenge von $\pi[f_0] \mid {}'S_0'$ ist aber mindestens eindimensional, wie sich oben ergeben hat.

2) Es werde nun die Annahme über die Normalität von X fallengelassen und nur noch vorausgesetzt, dass X in den Punkten von A irreduzibel ist.—

Sei $(''X, \nu, X)$ eine Normalisierung von X. Dann wird A vermöge ν^{-1} auf die lokal-analytische Menge $''A := \nu^{-1}(A) \subset ''X$ bijektiv bezogen; es folgt, dass $''A$ irreduzibel und dass die durch Beschränkung von ν bestimmte holomorphe Abbildung $''A \to A$ ein Homöomorphismus ist. Wird weiter $''f := f \triangle \nu$ $''X \xrightarrow[m]{} Y$ und $''f_0 := f_0 \triangle \nu : ''X \xrightarrow[m]{} Y_0$ gesetzt, so ist $f = ''f \triangle \nu^{-1}$, $f_0 = ''f_0 \triangle \nu^{-1}$ und $''f_0$ von $''f$ abhängig. Nach Voraussetzung ist $f \underset{m}{|} A$ definiert, sei $A_1 \subset A$ eine inbezug auf $f \mid A : A \xrightarrow[k]{} Y$ ausgezeichnete offene Menge. Für $''x \in \nu^{-1}(A_1)$ ist $''f(''x) = f \triangle \nu(''x) \subset f(\nu(''x))$. Wegen $\nu(''x) \in A_1$ ist die Menge $f(\nu(''x))$ einpunktig, dies trifft also auch für $''f(''x)$ zu. Somit ist $\nu^{-1}(A_1) \subset ''A$ eine inbezug auf $''f \mid ''A : \xrightarrow[k]{} Y$ ausgezeichnete Menge, $''f \underset{m}{|} ''A$ ist demnach definiert. Nach dem ersten Teil des Beweises ist nun $''f_0 \underset{m}{|} ''A$ definiert. Sei $''A_2 \subset ''A$ eine inbezug auf $''f_0 \mid ''A : ''A \xrightarrow[k]{} Y_0$ ausgezeichnete offene Menge. Für $x \in \nu(''A_2)$ ist

$$f_0(x) = ''f_0 \triangle \nu^{-1}(x) \subset ''f_0(\nu^{-1}(x)),$$

mithin ist $f_0(x)$ eine einpunktige Menge. $\nu(''A_2) \subset A$ ist daher eine inbezug auf $f_0 \mid A : A \xrightarrow[k]{} Y_0$ ausgezeichnete Menge, $f_0 \underset{m}{|} A$ ist also definiert.

Bemerkung. Es lassen sich leicht Beispiele dafür angeben, dass die Behauptung von Proposition 5.2.2 ohne die Voraussetzung der Irreduzibilität von X in den Punkten von A falsch ist.

PROPOSITION 5.2.3. *Sei $f : X \xrightarrow[m]{} Y$ eine meromorphe Abbildung, A eine irreduzible analytische Menge in X, derart, dass $f \underset{m}{|} A : A \xrightarrow[m]{} Y$ definiert und eigentlich und dass $r(f) = r(f \underset{m}{|} A)$ ist. Es sei $f_0 : X \xrightarrow[m]{} Y_0$ eine von f abhängige meromorphe Abbildung, für welche $f_0 \underset{m}{|} A$ definiert ist. Dann gilt:*

1) *$f_0 \underset{m}{|} A$ ist von $f \underset{m}{|} A$ abhängig;*

2) $(f_0 \underset{m}{|} A)(A) = f_0(A) = f_0(X)$;

3) $r(f_0) = r(f_0 \underset{m}{|} A)$.

Beweis. Wegen Proposition 4.2.2 lässt sich eine offene Menge $A_0 \subset A$ wählen, die sowohl inbezug auf $f \mid A : A \xrightarrow[k]{} Y$ wie auf $f_0 \mid A : A \xrightarrow[k]{} Y_0$ ausgezeichnet ist. A_0 ist dann auch inbezug auf $(f, f_0) \mid A$ ausgezeichnet. Für $x_0 \in A_0$ gilt $\{f, f_0\}(x_0) \subset (f, f_0)(x_0) = ((f, f_0) \mid A)(x_0)$ und

$$\{f \underset{m}{|} A, f_0 \underset{m}{|} A\}(x_0) \subset (f, f_0)(x_0) = ((f, f_0) \mid A)(x_0),$$

mithin sind $\{f, f_0\}(x_0)$ und $\{f \underset{m}{|} A, f_0 \underset{m}{|} A\}(x_0)$ einpunktige und gleiche Mengen. A_0 ist daher ausgezeichnet inbezug auf $\{f, f_0\} \,|\, A : A \xrightarrow[k]{} Y \times Y_0$ und es ist $\{f, f_0\} \,|\, A_0 = \{f \underset{m}{|} A, f_0 \underset{m}{|} A\} \,|\, A_0$; infolgedessen ist $\{f, f_0\} \underset{m}{|} A$ definiert und, wegen Proposition 3.4.2, mit $\{f \underset{m}{|} A, f_0 \underset{m}{|} A\}$ identisch. Es gilt

$$r(\{f, f_0\} \underset{m}{|} A) \leqq r(\{f, f_0\})$$

(dies ergibt sich durch Übergang zu der $\{f, f_0\}$ durch Liftung nach $G[\{f, f_0\}]$ zugeordneten Abbildung mit Proposition 1.2.1), daher ist

$$r(\{f \underset{m}{|} A, f_0 \underset{m}{|} A\} = r(\{f, f_0\} \underset{m}{|} A) \leqq r(\{f, f_0\}) = r(f) = r(f \underset{m}{|} A).$$

Andererseits ist $r(f \underset{m}{|} A) \leqq r(\{f \underset{m}{|} A, f_0 \underset{m}{|} A\})$, sodass $r(\{f \underset{m}{|} A, f_0 \underset{m}{|} A \ ) = r(f \underset{m}{|} A)$ folgt; $f_0 \underset{m}{|} A$ ist also von $f \underset{m}{|} A$ abhängig.

Es folgt auch $r(\{f, f_0\} \underset{m}{|} A) = r(\{f, f_0\})$. Weiter ist nach Proposition 4.3.1 mit $f \underset{m}{|} A$ auch $\{f \underset{m}{|} A, f_0 \underset{m}{|} A\} = \{f, f_0\} \underset{m}{|} A$ eigentlich. Nach Proposition 5.2.1 ist dann

$$\{f \underset{m}{|} A, f_0 \underset{m}{|} A\}(A) = (\{f, f_0\} \underset{m}{|} A)(A) = \{f, f_0\}(A) = \{f, f_0\}(X)$$

und $\{f, f_0\}(X) =: M$ eine irreduzible analytische Menge der Dimension $r(\{f, f_0\} \underset{m}{|} A) = r(f)$ in $Y \times Y_0$. Bezeichnet $\sigma_0 : Y \times Y_0 \to Y_0$ die natürliche Projektion, so ergibt sich mit Proposition 3.4.2, dass $f_0 = \sigma_0 \triangle \{f, f_0\}$ und $f_0 \underset{m}{|} A = \sigma_0 \triangle \{f \underset{m}{|} A, f_0 \underset{m}{|} A\}$ ist. Man hat demnach $f_0(X) = (\sigma_0 \triangle \{f, f_0\})(X) = (\sigma_0 \circ \{f, f_0\})(X) = \sigma_0(\{f, f_0\}(X)) = \sigma_0(M)$ und

$$\begin{aligned}(f_0 \underset{m}{|} A)(A) &= (\sigma_0 \triangle \{f \underset{m}{|} A, f_0 \underset{m}{|} A\})(A) \\ &= (\sigma_0 \circ \{f \underset{m}{|} A, f_0 \underset{m}{|} A\})(A) = \sigma_0(\{f \underset{m}{|} A, f_0 \underset{m}{|} A\}(A)) = \sigma_0(M),\end{aligned}$$

also $(f_0 \underset{m}{|} A)(A) = f_0(X)$ und dann auch $f_0(A) = f_0(X)$.

Es ist weiter $r(f_0) = r(\sigma_0 \,|\, M)$ und $r(f_0 \underset{m}{|} A) = r(\sigma_0 \,|\, M)$. Zum Nachweis der ersten dieser Beziehungen sei U eine offene nichtleere Menge mit abzählbarer Topologie in X, derart, dass jeder Punkt von U nichtsingulärer Punkt inbezug auf f_0 und $\{f, f_0\}$ ist und dass in jedem Punkt von U der lokale Rang der holomorphen Abbildung $\{f, f_0\} \,|\, U$ mit dem Rang $r(\{f, f_0\} \,|\, U) = r(\{f, f_0\}) = r(f)$ übereinstimmt. Es ist dann $\{f, f_0\}(U)$ eine offene Menge mit abzählbarer Topologie in M und $\sigma_0(\{f, f_0\}(U)) = f_0(U)$ eine Menge der topologischen Dimension $2r(\sigma_0 \mid M)$ in Y_0, andererseits ist die topologische Dimension von $f_0(U)$ gleich $2r(f_0)$ (vgl. [26], Satz 1). Demnach ist $r(f_0) = r(\sigma_0 \mid M)$. Entsprechend ist $r(f_0 \underset{m}{|} A) = r(\sigma_0 \mid M)$ nachzuweisen, sodass dann $r(f_0) = r(f_0 \underset{m}{|} A)$ folgt, wie unter 3) behauptet.

Bemerkung. Die Behauptung 1) der Proposition 5.2.3 gilt, wie sich aus dem Beweis ergibt, schon dann, wenn A nur eine lokal-analytische Menge ist und die Eigentlichkeit von $f \underset{m}{|} A$ nicht gefordert wird. Es lässt sich zeigen, dass 3) ebenfalls schon in diesem Falle zutrifft.

5.3. Wir sagen, die meromorphe Abbildung $f_0 \colon X \xrightarrow[m]{} Y_0$ *m-majorisiere* die meromorphe Abbildung $f_1 \colon X \xrightarrow[m]{} Y_1$, wenn eine meromorphe Abbildung $\alpha \colon Y_0 \xrightarrow[m]{} Y_1$ existiert, derart, dass $\alpha \triangle f_0$ definiert und $f_1 = \alpha \triangle f_0$ ist. Die meromorphe Abbildung $f^* \colon X \xrightarrow[m]{} Y^*$ heisst *m-maximal*, wenn sie surjektiv ist und wenn sie jede von f^* abhängige meromorphe Abbildung m-majorisiert. Ist f^* m-maximal und $f \colon X \xrightarrow[m]{} Y$ irgendeine mit f^* analytisch verwandte meromorphe Abbildung, so nennen wir das Paar (f^*, Y^*) eine *komplexe m-Basis zur Abbildung f.*

Es ist zweckmässig, noch eine Erweiterung des Begriffs der m-Maximalität einzuführen. Sei $f \colon X \xrightarrow[m]{} Y$ eine meromorphe Abbildung. Eine Klasse $\mathfrak{F}$ von meromorphen Abbildung von X heisst eine *f-Klasse*, wenn gilt:

(1) f gehört zu $\mathfrak{F}$.

(2) Jede Abbildung $\phi \in \mathfrak{F}$ ist von f abhängig.

(3) Gehören $\phi_1 \colon X \xrightarrow[m]{} Y_1$ und $\phi_2 \colon X \xrightarrow[m]{} Y_2$ zu $\mathfrak{F}$, so auch $\{\phi_1, \phi_2\} \colon X \xrightarrow[m]{} Y_1 \times Y_2$.

(4) Gehört $\phi \colon X \xrightarrow[m]{} Y'$ zu $\mathfrak{F}$ und ist $\phi(X)$ in einem komplexen Unterraum Y'_0 von Y' enthalten, so gehört die Beschränkung $\phi_0 \colon X \xrightarrow[m]{} Y'_0$ von ϕ auf Y'_0 zu $\mathfrak{F}$.

Eine surjektive meromorphe Abbildung f^* aus der f-Klasse $\mathfrak{F}$ heisse *$\mathfrak{F}$-maximal*, wenn jede Abbildung aus $\mathfrak{F}$ von f^* m-majorisiert wird.

Sind $f^* \colon X \xrightarrow[m]{} Y^*$, $f_1^* \colon X \xrightarrow[m]{} Y_1^*$ zwei $\mathfrak{F}$-maximale Abbildungen, so hat man surjektive meromorphe Abbildungen $\beta: Y_1^* \xrightarrow[m]{} Y^*$ und $\beta_1: Y^* \xrightarrow[m]{} Y_1^*$, derart, dass $f^* = \beta \triangle f_1^*$ und $f_1^* = \beta_1 \triangle f^*$ gilt. Wegen 5.1, Aussage (c), sind f^* und f_1^* demnach analytisch verwandt. Weiter ist $f^* = \beta \triangle (\beta_1 \triangle f^*) = (\beta \triangle \beta_1) \triangle f^*$; es folgt, dass die meromorphe Abbildung $\beta \triangle \beta_1: Y^* \xrightarrow[m]{} Y^*$ sicher in jedem Punkt der Menge

$$f^*(X - S(f^*)) - S(\beta \triangle \beta_1) \subset Y^*$$

mit der identischen Abbildung von Y^* übereinstimmt. Wegen der Surjektivität von f^* liegt $f^*(X - S(f^*))$ dicht in Y^*; mit Hilfe von Proposition 3.4.2 ergibt sich also, dass $\beta \triangle \beta_1$ die identische Abbildung von Y^* ist. Dann aber sind β und β_1 bimeromorph, und es ist $\beta_1 = \beta^{-1}$ (vgl. 5.1, p. 845). Wir nennen f^* und f_1^* *bimeromorph äquivalent.*

Zum Begriff der f-Klasse seien einige Beispiele betrachtet.

1) Die Klasse aller von der Abbildung $f: X \xrightarrow[m]{} Y$ abhängigen meromorphen Abbildung ist eine f-Klasse.

2) X sei eine irreduzible analytische Menge in einem komplexen Raum ${}^*X$, der in allen Punkten von X irreduzibel sei; ${}^*f: {}^*X \xrightarrow[m]{} Y$ sei eine meromorphe Abbildung, derart, dass ${}^*f \underset{m}{|} X =: f$ definiert und eigentlich und dass $r({}^*f) = r(f)$ ist. Es sei ${}^*\mathfrak{F}$ eine ${}^*f$-Klasse. Für jede Abbildung ${}^*\phi \in {}^*\mathfrak{F}$ existiert auf Grund von Proposition 5.2.2 die meromorphe Beschränkung ${}^*\phi \underset{m}{|} X =: \phi$. Es wird behauptet, dass diese Abbildungen ϕ eine f-Klasse $\mathfrak{F}$ (die auch mit ${}^*\mathfrak{F} \underset{m}{|} X$ bezeichnet sei) bilden.

Es ist zu prüfen, ob die obigen Bedingungen (1)-(4) für $\mathfrak{F}$ erfüllt sind: (1) trifft trivialerweise zu; (2) gilt auf Grund von Proposition 5.2.3. Zu (3): Sind Abbildungen $\phi_1 \in \mathfrak{F}$, $\phi_2 \in \mathfrak{F}$ mit

$$\phi_1 = {}^*\phi_1 \underset{m}{|} X, \qquad \phi_2 = {}^*\phi_2 \underset{m}{|} X \qquad ({}^*\phi_1 \in {}^*\mathfrak{F}, \phi_2 \in {}^*\mathfrak{F})$$

gegeben, so gehört $\{{}^*\phi_1, {}^*\phi_2\}$ zu ${}^*\mathfrak{F}$ und $\{{}^*\phi_1, {}^*\phi_2\} \underset{m}{|} X$ ist definiert. Sei U eine offene Menge in X, die zugleich inbezug auf die holomorphen Korrespondenzen ${}^*\phi_1 | X$ und ${}^*\phi_2 | X$ ausgezeichnet ist. Für $x \in U$ gilt

$$(\{{}^*\phi_1, {}^*\phi_2\} \underset{m}{|} X)(x) = ({}^*\phi_1, {}^*\phi_2)(x) = \{{}^*\phi_1 \underset{m}{|} X, {}^*\phi_2 \underset{m}{|} X\}(x);$$

auf Grund von Proposition 3.4.2 ist also

$$\{{}^*\phi_1, {}^*\phi_2\} \underset{m}{|} X = \{{}^*\phi_1 \underset{m}{|} X, {}^*\phi_2 \underset{m}{|} X\} = \{\phi_1, \phi_2\},$$

demnach gilt (3) für $\mathfrak{F}$.—Zu (4): Sei $\mathfrak{F} \ni \phi: X \xrightarrow[m]{} {}'Y$ mit $\phi = {}^*\phi \underset{m}{|} X$,

${}^*\phi \in {}^*\mathfrak{F}$; es sei $\phi(X)$ in einem irreduziblen komplexen Unterraum $'Y_0$ von $'Y$ enthalten, sei $\phi_0\colon X \xrightarrow[m]{} {}'Y_0$ die Beschränkung von ϕ auf $'Y_0$. Nach Proposition 5.2.3 ist ${}^*\phi({}^*X) = \phi(X)$, also ${}^*\phi({}^*X) \subset {}'Y_0$. Die Beschränkung ${}^*\phi_0\colon {}^*X \xrightarrow[m]{} {}'Y_0$ von ${}^*\phi$ auf $'Y_0$ gehört zu ${}^*\mathfrak{F}$, und es ist offenbar $\phi_0 = {}^*\phi_0 \underset{m}{|} X$. Also ist (4) für $\mathfrak{F}$ erfüllt.

3) X, ${}^*X$, f, ${}^*f$ mögen die gleiche Bedeutung wie im Beispiel 2) haben; es ist dann leicht ersichtlich, dass X in ${}^*X$ ein Fundamentalsystem $\mathfrak{U}(X)$ irreduzibler offener Umgebungen besitzt. Es sei $\mathfrak{F}$ die Klasse der wie folgt beschaffenen meromorphen Abbildungen von X: Zu jeder derartigen Abbildung ϕ existiere eine Umgebung ${}^*U_\phi \in \mathfrak{U}(X)$ und eine meromorphe Abbildung ${}^*\phi$ von ${}^*U_\phi$, derart, dass ϕ meromorphe Beschränkung von ${}^*\phi$ auf X und dass ${}^*\phi$ von ${}^*f \mid {}^*U_\phi$ abhängig ist. Dies lässt sich auch so ausdrücken: Jedes $\phi \in \mathfrak{F}$ ist "meromorphe Beschränkung" eines von f abhängigen "meromorphen Abbildungskeims auf X." Es wird wieder behauptet, dass $\mathfrak{F}$ eine f-Klasse ist.—Der Beweis verläuft entssprechend wie zu 2) (an Stelle der Abbildungen aus ${}^*\mathfrak{F}$ hat man meromorphe Abbildungen zu nehmen, die jeweils in zu $\mathfrak{U}(X)$ gehörenden Umgebungen von X definiert und dort von ${}^*f$ abhängig sind).

Die folgende Aussage ist unmittelbar nachzuprüfen:

PROPOSITION 5.3.1. *Sei* $f\colon X \xrightarrow[m]{} Y$ *eine meromorphe Abbildung,* $\tau\colon {}'X \xrightarrow[m]{} X$ *eine bimeromorphe Abbildung,* $\mathfrak{F}$ *eine f-Klasse meromorpher Abbildungen. Es sei* $'f := f \triangle \tau$ *gesetzt. Dann bilden die Abbildungen* $'\phi := \phi \triangle \tau$, $\phi \in \mathfrak{F}$, *eine 'f-Klasse* $'\mathfrak{F}$, *und die durch* $\phi \rightarrow \phi \triangle \tau$ *gegebene Zuordnung* $\mathfrak{F} \rightarrow {}'\mathfrak{F}$ *ist bijektiv. Einer* $\mathfrak{F}$*-maximalen Abbildung entspricht vermöge dieser Zuordnung eine* $'\mathfrak{F}$*-maximale Abbildung.*

5.4. Wir beweisen eine weitere Hauptaussage dieser Arbeit.

SATZ II. *Sei* $f\colon X \xrightarrow[m]{} Y$ *eine eigentliche meromorphe Abbildung. Dann gibt es in jeder f-Klasse* $\mathfrak{F}$ *von meromorphen Abbildungen eine mit f analytisch verwandte* $\mathfrak{F}$*-maximale eigentliche Abbildung.*

Beweis. Es darf vorausgesetzt werden, dass der komplexe Raum X normal und die Abbildung f holomorph ist. Denn trifft dies von vornherein nicht zu, so wird eine Normalisierung $('X, \nu, G[f])$ des Graphen $G[f]$ betrachtet; dann ist $('X, \pi[f] \circ \nu, X)$ eine eigentliche Modifikation und $'f := f \triangle (\pi[f] \circ \nu) = \hat{f} \circ \nu\colon {}'X \longrightarrow Y$ eine eigentliche holomorphe Abbildung. Es wird nun X, f durch $'X$, $'f$ ersetzt und gemäss Proposition 5.3.1 von $\mathfrak{F}$ zu der aus den meromorphen Abbildungen $'\phi := \phi \triangle (\pi[f] \circ \nu)$, $\phi \in \mathfrak{F}$, bestehenden $'f$-Klasse

$'\mathfrak{F}$ übergegangen; es genügt dann, die Behauptung für $'\mathfrak{F}$ statt für $\mathfrak{F}$ nachzuweisen.—Wir kehren zu den alten Bezeichnungen X, f, $\mathfrak{F}$ an Stelle von $'X$, $'f$, $'\mathfrak{F}$ zurück.

Wegen der Eigentlichkeit von f ist das Bild $f(X) := Y_0$ eine irreduzible $r(f)$-dimensionale analytische Menge in Y. Die Beschränkung $f_0 : X \longrightarrow Y_0$ von f auf Y_0 ist surjektiv, eigentlich, mit f analytisch verwandt und sie gehört zu $\mathfrak{F}$. Wir betrachen die Klasse $\mathfrak{K}$ derjenigen surjektiven meromorphen Abbildungen aus $\mathfrak{F}$, die f_0 m-majorisieren (insbesondere gehört f_0 zu $\mathfrak{K}$). Ist $f_i : X \underset{m}{\longrightarrow} Y_i$ irgendeine Abbildung aus $\mathfrak{K}$, so ist Y_i irreduzibel, und man hat eine meromorphe Abbildung $\alpha_i : Y_i \underset{m}{\longrightarrow} Y_0$ mit $f_0 = \alpha_i \triangle f_i$; daher ist f_0 von f_i abhängig. Andererseits ist f_i von f, also von f_0 abhängig, infolgedessen sind f_i und f_0 analytisch verwandt, und es ist $r(f_i) = r(f_0)$. Nach Proposition 4.3.2 ist mit f_0 auch f_i eigentlich, und gleiches trifft, da f_i surjektiv ist, für α_i zu; ferner ist α_i mit f_0 surjektiv. Demnach gilt $r(\alpha_i) = \dim Y_i = r(f_i) = r(f_0) = \dim Y_0$. Das Tripel (Y_i, α_i, Y_0) ist also eine meromorphe Überlagerung von Y_0, sei $\mathfrak{g}_i := \mathfrak{g}(\alpha_i)$ ihre Blätterzahl. Es gibt jeweils in Y_i bezw. Y_0 eine von Y_i bezw. Y_0 verschiedene analytische Menge M_i bezw. M_{0i}, derart, dass $\alpha_i(M_i) = M_{0i}$ sowie $\alpha_i^{-1}(M_{0i}) = M_i$ ist, dass $\alpha_i \mid Y_i - M_i$ eine holomorphe Abbildung und dass die durch Beschränkung von α_i bestimmte holomorphe Abbildung $'\alpha_i : Y_i - M_i \rightarrow Y_0 - M_{0i}$ eigentlich und lokal biholomorph ist; über jedem Punkt von $Y_0 - M_{0i}$ liegen dann vermöge $'\alpha_i$ genau $\mathfrak{g}_i$ Punkte von $Y_i - M_i$.

Die Menge $S_i := f_i(S(f_i))$ ist eine analytische Menge in Y_i; wir behaupten, dass sie von Y_i verschieden ist. Angenommen, es sei $S_i = Y_i$. Der Rang der eigentlichen holomorphen Korrespondenz $f_i \mid S(f_i) : S(f_i) \underset{k}{\longrightarrow} Y_i$ ist dann gleich $\dim Y_i$, und es gibt eine irreduzible Komponente S_i' der analytischen Menge $\pi[f_i]^{-1}(S(f_i)) \subset G[f_i]$, derart, dass $r(\hat{f} \mid S_i') = \dim Y_i = r(f_i) = r(\hat{f}_i) = r(f_0)$ ist. Andererseits ist die Beschränkung von f_0 auf irgendeine irreduzible Komponente von $S(f_i)$ eine holomorphe Abbildung, deren Rang kleiner als $r(f_0)$ ist; denn sonst müsste nach Proposition 5.2.2 die meromorphe Beschränkung von f_i auf diese Komponente definiert sein, und dies würde wegen der Normalität von X bedeuten, dass auf der Komponente nichtsinguläre Punkte inbezug auf f_i lägen, was nicht zutrifft. Es folgt, dass auch $r(f_0 \circ \pi[f_i] \mid S_i') < r(f_0)$ sein muss. Nun ist die eigentliche holomorphe Abbildung $f_0 \circ \pi[f_i] : G[f_i] \rightarrow Y_0$ von der eigentlichen holomorphen Abbildung $\hat{f}_i = f_i \triangle \pi[f_i] : G[f_i] \rightarrow Y_i$ abhängig ($f_0 \circ \pi[f_i]$ und f_i sind sogar analytisch verwandt, weil f_0 und f_i es sind). Nach Proposition 5.2.3, Behauptung 3) gilt dann $r(f_0 \circ \pi[f_i] \mid S_i') = r(f_0 \circ \pi[f_i])$; da $r(f_0 \circ \pi[f_i]) = r(f_0)$

ist, folgt also $r(f_0 \circ \pi[f_i] \mid S_i') = r(f_0)$, womit ein Widerspruch hergestellt ist. S_i ist mithin in Y_i echt enthalten.

Sei jetzt $B_{0i} := Y_0 - (M_{0i} \cup \alpha_i(S_i))$ und $B_i := Y_i - (M_i \cup \alpha_i^{-1}(\alpha_i(S_i)))$; B_{0i} und B_i sind irreduzible offene dichte Unterräume von Y_0 bezw. Y_i, und es ist $\alpha_i(B_i) = {}'\alpha_i(B_i) = B_{0i}$ sowie $\alpha_i^{-1}(B_{0i}) = {}'\alpha_i^{-1}(B_{0i}) = B_i$. Es sei $''\alpha_i \colon B_i \to B_{0i}$ die durch Beschränkung von α_i bestimmte lokal-biholomorphe eigentliche Überlagerungsabbildung. Ist x_{0i} ein Punkt aus B_{0i} und bezeichnen $x_i^{(\gamma)}$ $(\gamma = 1, \cdots, \mathfrak{g}_i)$ die $\mathfrak{g}_i$ Urbildpunkte von x_{0i} vermöge $''\alpha_i$ in B_i, so hat man $f_0^{-1}(x_{0i}) = f_i^{-1}(''\alpha_i^{-1}(x_{0i})) = f_i^{-1}(\bigcup_{\gamma} \{x_i^{(\gamma)}\}) = \bigcup_{\gamma} f_i^{-1}(x_i^{(\gamma)})$. Da $x_i^{(\gamma)}$ nicht in S_i liegt, sind die $\mathfrak{g}_i$ Fasern $f_i^{-1}(x_i^{(\gamma)})$ paarweise punktfremd. Die Faser $f_0^{-1}(x_{0i})$ zerfällt daher in mindestens $\mathfrak{g}_i$ zusammenhängende Komponenten, und dies gilt für jeden Punkt $x_{0i} \in B_{0i}$. Andererseits sind wegen der Eigentlichkeit der holomorphen Abbildung f_0 alle Fasern von f_0 kompakte analytische Mengen, und diese zerfallen in jeweils endlich viele zusammenhängende Komponenten. Es gibt also (Proposition 1.2.3) eine positive ganze Zahl l_0, derart, dass die Menge derjenigen Punkte $y_0 \in Y_0$, für welche die Faser $f_0^{-1}(y_0)$ höchstens l_0 zusammenhängende Komponenten besitzt, in einer nichtleeren offenen Teilmenge E_0 von Y_0 dicht liegt. Da B_{0i} in Y_0 dicht ist, ist die offene Menge $B_{0i} \cap E_0$ nichtleer, demnach muss $\mathfrak{g}_i \leqq l_0$ gelten. Die den $f_i \in \mathfrak{R}$ zugeordneten Blätterzahlen bleiben also unter einer endlichen Schranke.

Es sei nun $f_j \colon X \xrightarrow[m]{} Y_j$ eine meromorphe Abbildung aus $\mathfrak{R}$, derart, dass die Blätterzahl $\mathfrak{g}_j$ der zugehörigen meromorphen Überlagerung (Y_j, α_j, Y_0) maximal ist. Wir behaupten, dass f_j eine gesuchte $\mathfrak{F}$-maximale Abbildung ist. In der Tat, ist $\phi \colon X \xrightarrow[m]{} Y^*$ irgendeine Abbildung aus $\mathfrak{F}$, so gehört auch $\{f_j, \phi\} \colon X \xrightarrow[m]{} Y_j \times Y^*$ zu $\mathfrak{F}$. f_j ist mit f_0, also mit f, analytisch verwandt; da ϕ von f abhängig ist, ist ϕ also auch von f_j abhängig, d.h. es gilt $r(f) = r(f_j) = r(\{f_j, \phi\})$. Nach Proposition 4.3.1 ist weiter $\{f_j, \phi\}$ mit f_j eigentlich. Daher ist das Bild $\{f_j, \phi\}(X) =: Y_0^*$ eine irreduzible $r(f)$-dimensionale analytische Menge in $Y_j \times Y^*$. Die Beschränkung $\phi_0^* \colon X \xrightarrow[m]{} Y_0^*$ von $\{f_j, \phi\}$ auf Y_0^* gehört zu $\mathfrak{F}$, und sie ist surjektiv, eigentlich und mit $\{f_j, \phi\}$, also mit f und f_0, analytisch verwandt. Es bezeichne $\epsilon_0 \colon Y_0^* \to Y_j \times Y^*$ die Injektion, weiter σ_j bezw. σ^* die Projektion von $Y_j \times Y^*$ auf Y_j bezw. auf Y^*. Man hat $f_j = \sigma_j \circ \epsilon_0 \circ \phi_0^*$, demnach wird f_j von ϕ_0^* m-majorisiert. Es ergibt sich wiederum, dass $(Y_0^*, \sigma_j \circ \epsilon_0, Y_j)$ eine m-Überlagerung ist ($\sigma_j \circ \epsilon_0$ ist sogar eine holomorphe Abbildung). Dann ist auch $(Y_0^*, \alpha_j \triangle (\sigma_j \circ \epsilon_0), Y_0)$ eine m-Überlagerung, und es ist

$$\mathfrak{g}(\alpha_j \triangle (\sigma_j \circ \epsilon_0)) = \mathfrak{g}(\alpha_j) \cdot \mathfrak{g}(\sigma_j \circ \epsilon_0) = \mathfrak{g}_j \cdot \mathfrak{g}(\sigma_j \circ \epsilon_0).$$

Nun gilt $f_0 = \alpha_j \triangle f_j = \alpha_j \triangle ((\sigma_j \circ \epsilon_0) \circ \phi_0^*) = (\alpha_j \triangle (\sigma_j \circ \epsilon_0)) \triangle \phi_0^*$, also wird f_0 von ϕ_0^* m-majorisiert; mithin gehört ϕ_0^* zu $\mathfrak{K}$, und $(Y_0^*, \alpha_j \triangle (\sigma_j \circ \epsilon_0), Y_0)$ ist die zu ϕ_0^* gehörige m-Überlagerung von Y_0. Wegen der Maximalität von $\mathfrak{g}_j$ muss $\mathfrak{g}(\alpha_j \triangle (\sigma_j \circ \epsilon_0)) = \mathfrak{g}_j$ gelten; daher ist $\mathfrak{g}(\sigma_j \circ \epsilon_0) = 1$, und es folgt, dass $\sigma_j \circ \epsilon_0$ eine bimeromorphe Abbildung ist. Es ist dann

$$\begin{aligned}(\sigma_j \circ \epsilon_0)^{-1} \triangle f_j &= (\sigma_j \circ \epsilon_0)^{-1} \triangle ((\sigma_j \circ \epsilon_0) \circ \phi_0^*)\\ &= ((\sigma_j \circ \epsilon_0)^{-1} \triangle (\sigma_j \circ \epsilon_0)) \triangle \phi_0^* = \phi_0^*.\end{aligned}$$

Nun gilt $\phi = \sigma^{\ddagger} \circ \epsilon_0 \circ \phi_0^*$, demnach hat man $\phi = \sigma^* \circ \epsilon_0 \circ ((\sigma_j \circ \epsilon_0)^{-1} \triangle f_j) = ((\sigma^* \circ \epsilon_0) \circ (\sigma_j \circ \epsilon_0)^{-1}) \triangle f_j$. Also wird ϕ von f_j m-majorisiert, w. z. b. w.

Spezialisierung auf den Fall der f-Klasse aller von f abhängigen meromorphen Abbildungen ergibt das

KOROLLAR ZU SATZ II. *Zu jeder eigentlichen meromorphen Abbildung* $f: X \xrightarrow[m]{} Y$ *existiert eine mit* f *analytisch verwandte m-maximale eigentliche meromorphe Abbildung.*

5. 5. Satz II lässt sich wie folgt erweitern:

SATZ III. *Sei* $f: X \xrightarrow[m]{} Y$ *eine meromorphe Abbildung, ferner* A *eine irreduzible analytische Menge in* X*, derart, dass die holomorphe Korrespondenz* $a: = f \mid A: A \xrightarrow[k]{} Y$ *wenigstens eine irreduzible Komponente* $a': A \xrightarrow[k]{} Y$ *besitzt, die eigentlich ist und für die* $r(a') = r(f)$ *gilt. Es sei* $\mathfrak{F}$ *eine f-Klasse meromorpher Abbildungen. Dann existiert eine mit* f *analytisch verwandte* $\mathfrak{F}$*-maximale meromorphe Abbildung* $f^*: X \xrightarrow[m]{} Y^*$ *auf einen* $r(f)$*-dimensionalen komplexen Raum* Y^*.

Beweis. Sei $('X, \nu, G[f])$ eine Normalisierung des Graphen $G[f]$, es sei $\tau: = \pi[f] \circ \nu$ gesetzt; dann ist $('X, \tau, X)$ eine eigentliche Modifikation und $'f: = f \triangle \tau = \hat{f} \circ \nu: 'X \to Y$ eine holomorphe Abbildung. Auf Grund der Voraussetzung über f besitzt die analytische Menge $\tau^{-1}(A)$ eine irreduzible Komponente $'A$, derart, dass $'f \mid 'A: 'A \to Y$ eigentlich und $r('f \mid 'A) = r('f) = r(f)$ ist. Der f-Klasse $\mathfrak{F}$ entspricht nach Proposition 5. 3. 1 vermöge der Zuordnung $\phi \to \phi \triangle \tau$, $\phi \in \mathfrak{F}$, bijektiv eine $'f$-Klasse $'\mathfrak{F}$. Für jede Abbildung $'\phi \in '\mathfrak{F}$ ist nach Proposition 5. 2. 2 die meromorphe Beschränkung $'\phi \underset{m}{|} 'A$ definiert, und nach Proposition 5. 2. 3 ist $'\phi('X) = '\phi('A) = ('\phi \underset{m}{|} 'A)('A)$ und $r('\phi) = r('\phi \underset{m}{|} 'A)$. Es folgt, dass für $\phi = '\phi \triangle \tau^{-1} \in \mathfrak{F}$ die holomorphe Korrespondenz $\alpha: = \phi \mid A$ stets eine irreduzible Komponente α' besitzt, derart, dass $\phi(X) = \phi(A) = \alpha(A) = \alpha'(A)$ und $r(\alpha') = r(\phi)$ gilt; ist $'\phi \underset{m}{|} 'A$ speziell eigentlich, so kann auch α' als eigentlich gewählt werden, und die Bildmenge

$\alpha'(A) = \phi(X)$ ist dann eine irreduzible $r(\phi)$-dimensionale analytische Menge. —Auf Grund von Proposition 5.3.1 genügt es, die Existenz einer mit $'f$ analytisch verwandten $'\mathfrak{F}$-maximalen meromorphen Abbildung auf einen $r('f)$-dimensionalen komplexen Raum nachzuweisen. Wir benutzen im folgenden statt $'X$, $'A$, $'f$, $'\mathfrak{F}$ der Einfachheit halber wieder die Bezeichnungen X, A, f, $\mathfrak{F}$.

Es sei $f' := f \mid A$ gesetzt; dann bilden die meromorphen Beschränkungen $\phi' := \phi \underset{m}{|} A$, $\phi \in \mathfrak{F}$, eine f'-Klasse $\mathfrak{F}' = \mathfrak{F} \mid A$ (vgl. oben Beispiel 2), p. 851). Nach Satz II existiert eine mit f' analytisch verwandte $\mathfrak{F}'$-maximale eigentliche meromorphe Abbildung $f'^{*}: A \xrightarrow[m]{} Y^{*}$; wegen $r(f'^{*}) = r(f')$ und $r(f') = r(f)$ ist Y^{*} $r(f)$-dimensional. Sei $f^{*}: X \xrightarrow[m]{} Y^{*}$ eine meromorphe Abbildung aus $\mathfrak{F}$, derart, dass $f'^{*} = f^{*} \underset{m}{|} A$ ist. Wir behaupten, dass f^{*} eine gesuchte $\mathfrak{F}$-maximale Abbildung ist.

Sicher ist f^{*} surjektiv, da dies für f'^{*} zutrifft. Sei jetzt irgendeine Abbildung $\phi: X \xrightarrow[m]{} 'Y$ aus $\mathfrak{F}$ gegeben. Zur meromorphen Beschränkung $\phi' := \phi \underset{m}{|} A : A \xrightarrow[m]{} 'Y$ existiert dann eine meromorphe Abbildung μ: $Y^{*} \xrightarrow[m]{} 'Y$ mit $\phi' = \mu \bigtriangleup f'^{*}$. Wir werden zeigen, dass auch $\phi = \mu \bigtriangleup f^{*}$ gilt, womit die $\mathfrak{F}$-Maximalität von f^{*} nachgewiesen sein wird; da insbesondere dann f durch f^{*} m-majorisiert wird, folgt auch, dass f^{*} mit f analytisch verwandt ist.

Die meromorphe Abbildung $\tilde{\phi} := \mu \bigtriangleup f^{*}: X \xrightarrow[m]{} 'Y$ ist von f^{*}, also von f, abhängig; nach Proposition 5.2.2 existiert die meromorphe Beschränkung $\tilde{\phi} \underset{m}{|} A$. Sei $U \subset A$ eine zugleich inbezug auf die holomorphen Korrespondenzen $\tilde{\phi} \mid A$, $f^{*} \mid A$, $\mu \circ f'^{*}$ ausgezeichnete offene Menge, dann gilt für $x \in U$:

$$\begin{aligned}(\tilde{\phi} \underset{m}{|} A)(x) &= (\tilde{\phi} \mid A)(x) = \tilde{\phi}(x) = (\mu \bigtriangleup f^{*})(x) \subset (\mu \circ f^{*})(x) \\ &= \mu(f^{*}(x)) = \mu((f^{*} \mid A)(x)) = \mu((f^{*} \underset{m}{|} A)(x)) \\ (\mu \circ f'^{*})(x) &= (\mu \bigtriangleup f'^{*})(x) = \phi'(x),\end{aligned}$$

also $(\tilde{\phi} \underset{m}{|} A)(x) = \phi'(x)$, da $(\mu \circ f'^{*})(x) = \phi'(x)$ einpunktig ist. Nach Proposition 3.4.2 ist daher $\tilde{\phi} \underset{m}{|} A = \phi' = \phi \underset{m}{|} A$. Wir behaupten nun allgemein: *Sind* $\phi: X \xrightarrow[m]{} 'Y$, $\tilde{\phi}: X \xrightarrow[m]{} 'Y$ *irgend zwei von* f *abhängige meromorphe Abbildungen, für welche* $\phi \underset{m}{|} A = \tilde{\phi} \underset{m}{|} A$ *gilt, so ist* $\phi = \tilde{\phi}$.

Um dies nachzuweisen, wird die Abbildung $\{\phi, \tilde{\phi}\}: X \xrightarrow[m]{} 'Y \times 'Y$

betrachtet; $\{\phi, \bar{\phi}\}$ ist von f abhängig. Auf Grund von Proposition 5.2.2 sind die meromorphen Beschränkungen $\phi \underset{m}{|} A$, $\bar{\phi} \underset{m}{|} A$, $\{\phi, \bar{\phi}\} \underset{m}{|} A$ definiert. Ist $V \subset \dot{A}$ eine zugleich inbezug auf die holomorphen Korrespondenzen $\phi \mid A$, $\bar{\phi} \mid A$, $\{\phi, \bar{\phi}\} \mid A$ ausgezeichnete offene Menge, so gilt für $x \in V$:

$$(\{\phi, \bar{\phi}\} \underset{m}{|} A)(x) = (\{\phi, \bar{\phi}\} \mid A)(x) = \{\phi, \bar{\phi}\}(x) \subset (\phi, \bar{\phi})(x)$$

$$= (\phi \mid A, \bar{\phi} \mid A)(x) = (\phi \underset{m}{|} A, \bar{\phi} \underset{m}{|} A)(x) = \{\phi \underset{m}{|} A, \bar{\phi} \underset{m}{|} A\}(x),$$

also $(\{\phi, \bar{\phi}\} \underset{m}{|} A)(x) = \{\phi \underset{m}{|} A, \bar{\phi} \underset{m}{|} A\}(x)$; nach $\{\phi, \bar{\phi}\} \underset{m}{|} A = \{\phi \underset{m}{|} A, \bar{\phi} \underset{m}{|} A\}$. Nach Proposition 5.2.3 gilt $\{\phi, \bar{\phi}\}(X) = (\{\phi, \bar{\phi}\} \underset{m}{|} A)(A)$, somit ist $\{\phi, \bar{\phi}\}(X) = \{\phi \underset{m}{|} A, \bar{\phi} \underset{m}{|} A\}(A)$. Wegen $\phi \underset{m}{|} A = \bar{\phi} \underset{m}{|} A$ ist nun $\{\phi \underset{m}{|} A, \bar{\phi} \underset{m}{|} A\}(A)$ in der Diagonale von $'Y \times {}'Y$ enthalten. Gleiches trifft also für $\{\phi, \bar{\phi}\}(X)$ zu, woraus folgt, dass ϕ und $\bar{\phi}$ jedenfalls in allen Punkten von X, die zugleich inbezug auf ϕ und $\bar{\phi}$ ausgezeichnet sind, übereinstimmen. Auf Grund von Proposition 3.4.2 muss dann $\phi = \bar{\phi}$ gelten, wie behauptet.—Damit ist auch Satz III bewiesen.

Wir notieren noch die folgenden

ZUSÄTZE ZU SATZ III. (1) *Die holomorphe Korrespondenz* $a^* := f^* \mid A : A \xrightarrow[k]{} Y^*$ *besitzt eine irreduzible Komponente* $a^{*\prime}: A \xrightarrow[k]{} Y^*$, *die eigentlich ist und für die* $r(a^{*\prime}) = r(f^*) = r(f)$ *sowie* $a^{*\prime}(A) = Y^*$ *gilt.*

(2) *Die meromorphe Abbildung* $\beta: Y^* \xrightarrow[m]{} Y$, *für welche* $f = \beta \triangle f^*$ *gilt, ist eigentlich und vom Range* $r(\beta) = r(f)$.

(3) *Ist* ϕ *irgendeine meromorphe Abbildung aus* $\mathfrak{Y}$, *so besitzt die holomorphe Korrespondenz* $\alpha := \phi \mid A$ *wenigstens eine irreduzible Komponente* α', *derart, dass* $r(\alpha') = r(\phi)$ *und* $\alpha'(A) = \phi(A) = \phi(X)$ *ist.*

Beweis. Die Behauptungen (1) und (3) ergeben sich aus dem obigen Beweis zu Satz III.—Zu (2): Die Behauptung trifft jedenfalls zu, wenn der Raum X normal und die Abbildung f holomorph ist. Denn es ist dann $f^* \underset{m}{|} A : A \xrightarrow[m]{} Y^*$ eine eigentliche surjektive meromorphe Abbildung (vgl. den 2. Absatz des Beweises zu Satz III); wegen $f \mid A = \beta \triangle (f^* \underset{m}{|} A)$ und der Eigentlichkeit von $f \mid A$ ist β nach Proposition 4.3.2 also eigentlich. Die Bildmenge $\beta(Y^*)$ ist mithin eine analytische Menge der Dimension $r(\beta)$ in Y; da die $r(f)$-dimensionale analytische Menge $(f \mid A)(A) = f(X)$ in $\beta(Y^*)$ enthalten ist, gilt $r(f) \leqq r(\beta)$; wegen $r(\beta) \leqq \dim Y^* = r(f)$ ist daher $r(\beta) = r(f)$. Auf diesen Sonderfall lässt sich der Allgemeinfall mittels einer eigentlichen Modifikation $('X, \tau, X)$, derart, dass der komplexe Raum $'X$ normal und die Abbildung $'f := f \triangle \tau$ holomorph ist zurückführen (vgl. den 1. Absatz des Beweises zu Satz III): Der f-Klasse $\mathfrak{F}$ entspricht vermöge τ eine $'f$-Klasse $'\mathfrak{F}$, und die Abbildung $'f^* := f^* \triangle \tau$ ist $'\mathfrak{F}$-maximal; wegen

$$'f = f \triangle \tau = (\beta \triangle f^*) \triangle \tau = \beta \triangle (f^* \triangle \tau) = \beta \triangle 'f^*$$

folgt also die Behauptung.

Bemerkung. Es ist leicht zu zeigen, dass die in Satz III und den Zusätzen behaupteten Aussagen über f^* nicht nur für eine spezielle sondern für *jede* $\mathfrak{F}$-maximale meromorphe Abbildung zutreffen.

Eine Satz III entsprechende Aussage gilt auch für "meromorphe Abbildungskeime" auf der analytischen Menge $A \subset X$ (die jeweils durch meromorphe Abbildungen von Umgebungen von A repräsentiert werden; vgl. oben Beispiel 3), p. 852). Wir begnügen uns damit, die Aussage hier nur anzugeben; der Beweis lässt sich, mutatis mutandis, wie der Beweis zu Satz III führen. Ist $\bar{f}$ ein meromorpher Abbildungskeim auf A, so ist naheliegend, wie die Begriffe "$\bar{f}$-Klasse $\bar{\mathfrak{F}}$ von meromorphen Abbildungskeimen" und "$\bar{\mathfrak{F}}$-maximaler meromorpher Abbildungskeim" zu definieren sind.

Satz III′. *Sei* $f: X \underset{m}{\longrightarrow} Y$ *eine meromorphe Abbildung, ferner* A *eine irreduzible analytische Menge in* X, *die ein Fundamentalsystem irreduzibler offener Umgebungen in* X *besitzt. Die holomorphe Korrespondenz* $a := \hat{f} \,|\, A$: $A \underset{k}{\longrightarrow} Y$ *besitze wenigstens eine irreduzible Komponente* $a' : A \underset{k}{\longrightarrow} Y$, *die eigentlich ist und für die* $r(a') = r(f)$ *gilt. Es sei* $\bar{f}$ *der durch* f *repräsentierte meromorphe Abbildungskeim auf* A *und* $\bar{\mathfrak{F}}$ *eine* $\bar{f}$-*Klasse meromorpher Abbildungskeime. Dann existiert ein* $\bar{\mathfrak{F}}$-*maximaler meromorpher Abbildungskeim.*

Bemerkung. Die Forderung, dass die analytische Menge A ein Fundamentalsystem irreduzibler offener Umgebungen besitzen soll, dient—ebenso wie die Voraussetzung der Irreduzibilität der Urbildräume meromorpher Abbildungen—lediglich zur Vereinfachung der Formulierung.—Es ist klar, dass sich Satz III leicht aus Satz III′ folgern lässt.

5.6. Wir wollen den Spezialfall von Satz III bezw. Satz III′, in welchem die analytische Menge A nulldimensional ist, also aus genau einem Punkt besteht, besonders hervorheben.

Ein Punkt $x_0 \in X$ heisst eine *Unbestimmtheitsstelle k. Stufe* der meromorphen Abbildung $f: X \underset{m}{\longrightarrow} Y$, wenn $\dim f(x_0) = k$ ist (nach 4.1 ist $f(x_0)$ eine kompakte analytische Menge in Y; vermöge $\hat{f}: G[f] \rightarrow Y$ ist $\pi[f]^{-1}(x_0)$

biholomorph auf $f(x_0)$ bezogen). x_0 heisst *Unbestimmtheitsstelle maximaler Stufe* von f, wenn $\dim f(x_0) = r(f)$ ist.

Da für einen Punkt $x_0 \in X$ stets $f \mid \{x_0\} : \{x_0\} \xrightarrow[k]{} Y$ eine eigentliche holomorphe Korrespondenz und $r(f \mid \{x_0\}) = \dim f(x_0)$ ist, folgt aus dem Zusatz (3) zu Satz III: Ist x_0 eine Unbestimmtheitsstelle maximaler Stufe von f, so ist $f(x_0) = f(X)$ (die analytische Menge $f(x_0)$ ist also insbesondere irreduzibel). Ist weiter $f_0 : X \xrightarrow[m]{} Y_0$ eine von f abhängige meromorphe Abbildung, so ist x_0 auch Unbestimmtheitsstelle maximaler Stufe von f_0.

Aus Satz III bezw. Satz III′ ergibt sich nun

Satz IIIa. *Sei $f : X \xrightarrow[m]{} Y$ eine meromorphe Abbildung mit einer Unbestimmtheitsstelle x_0 maximaler Stufe. Es sei $\mathfrak{F}$ eine f-Klasse meromorpher Abbildungen. Dann existiert eine mit f analytisch verwandte $\mathfrak{F}$-maximale meromorphe Abbildung $f^* : X \xrightarrow[m]{} Y^*$ auf einen $r(f)$-dimensionalen kompakten komplexen Raum Y^*.*

Satz III′a. *Sei $f : X \xrightarrow[m]{} Y$ eine meromorphe Abbildung mit einer Unbestimmtheitsstelle x_0 maximaler Stufe; x_0 besitze ein Fundamentalsystem irreduzibler offener Umgebungen in X. Es sei $\tilde{f}$ der durch f repräsentierte meromorphe Abbildungskeim in x_0 und $\tilde{\mathfrak{F}}$ eine $\tilde{f}$-Klasse meromorpher Abbildungskeime. Dann existiert ein $\tilde{\mathfrak{F}}$-maximaler meromorpher Abbildungskeim.*

6. Abhängige meromorphe Abbildungen.

6.1. Ein System $\mathfrak{S}$ von l meromorphen Abbildungen $f_\lambda : X \xrightarrow[m]{} Y_\lambda$ $(\lambda = 1, \cdots, l)$ des irreduziblen komplexen Raumes X heisse *unabhängig,* wenn kein f_λ konstant ist und $r(\{f_1, \cdots, f_l\}) = \sum_{\lambda=1}^{l} r(f_\lambda)$ gilt (für $l = 1$ ist $\{f_1\} := f_1$, die Rangbedingung also trivial). Ist das System nicht unabhängig, so nennen wir es *schwach abhängig.* $\mathfrak{S}$ heisse *abhängig* im Falle $l = 1$, wenn f_1 konstant ist, und im Falle $l > 1$, wenn es eine Anordnung $f_{\lambda_1}, \cdots, f_{\lambda_l}$ der f_λ gibt, derart, dass f_{λ_l} von der meromorphen Abbildung $\{f_{\lambda_1}, \cdots, f_{\lambda_{l-1}}\}$ abhängig ist (es ist dann

$$r(\{f_1, \cdots, f_l\}) = r(\{f_{\lambda_1}, \cdots, f_{\lambda_l}\}) = r(\{f_{\lambda_1}, \cdots, f_{\lambda_{l-1}}\}) \leqq \sum_{i=1}^{l-1} r(f_{\lambda_i})).$$

Offenbar ist $\mathfrak{S}$ stets dann ein schwach abhängiges System, wenn es abhängig ist; das Umgekehrte gilt im allgemeinen nicht. Ist wenigstens eine der Abbildungen f_λ konstant, so ist $\mathfrak{S}$ ein abhängiges System.

Eine Bedingung für die Unabhängigkeit von $\mathfrak{S}$ liefert

PROPOSITION 6.1.1. *Ist keine der meromorphen Abbildungen* f_λ: $X \xrightarrow[m]{} Y_\lambda$ $(\lambda = 1, \cdots, l)$ *konstant und existiert eine lokal-analytische Menge* A *in* X, *derart, dass die holomorphe Korrespondenz* $a := \{f_1, \cdots, f_l\} \mid A$: $A \xrightarrow[k]{} Y_1 \times \cdots \times Y_l$ *den Rang* $\sum_{\lambda=1}^{l} r(f_\lambda)$ *hat, so bilden* $f_1, \cdots, f_l$ *ein unabhängiges System.*

Denn es ist stets $r(\{f_1, \cdots, f_l\}) \leqq \sum_{\lambda=1}^{l} r(f_\lambda)$, und auf Grund der Voraussetzung muss hier das Gleichheitszeichen gelten.

Trifft für das System $\mathfrak{S}$ der f_λ und die lokal-analytische Menge $A \subset X$ die Voraussetzung von Proposition 6.1.1 zu, so nennen wir $\mathfrak{S}$ *auf* A *stark unabhängig.* Besteht insbesondere A aus einem einzigen Punkt x_0, so heisst $\mathfrak{S}$ *in* x_0 *stark unabhängig;* x_0 ist dann notwendig eine Unbestimmtheitsstelle maximaler Stufe von $\{f_1, \cdots, f_l\}$, und das gleiche gilt nach 5.6 auch für jedes f_λ, da f_λ von $\{f_1, \cdots, f_l\}$ abhängig ist. Ist B irgendeine nichtleere offene Menge in X, so ist klar, dass $\mathfrak{S}$ genau dann ein unabhängiges System ist, wenn $\mathfrak{S}$ auf B stark unabhängig ist.

Das System $\mathfrak{S}$ der meromorphen Abbildungen $f_1, \cdots, f_l$ heisse weiter *$\mathfrak{B}$-unabhängig* (*"Bildraum-unabhängig"*), wenn kein f_λ konstant ist und das Bild von X unter der meromorphen Abbildung $\{f_1, \cdots, f_l\}$ nicht global analytisch dünn im Bildraum, d.h. dort in keiner analytischen Menge von positiver Codimension enthalten ist. Ist $\mathfrak{S}$ nicht $\mathfrak{B}$-unabhängig, so nennen wir $\mathfrak{S}$ *schwach $\mathfrak{B}$-abhängig.* Die meromorphe Abbildung f_0: $X \xrightarrow[m]{} Y_0$ heisse von der meromorphen Abbildung f: $X \xrightarrow[m]{} Y$ *$\mathfrak{B}$-abhängig,* wenn die Bildmenge $\{f, f_0\}(X)$ in einer analytischen Menge der Dimension $r(f)$ enthalten ist. Das System $\mathfrak{S}$ der Abbildungen $f_1, \cdots, f_l$ heisse schliesslich *$\mathfrak{B}$-abhängig* im Falle $l = 1$, wenn f_1 konstant ist, und im Falle $l > 1$, wenn es eine Anordnung $f_{\lambda_1}, \cdots, f_{\lambda_l}$ der f_λ gibt, derart, dass f_{λ_l} von $\{f_{\lambda_1}, \cdots, f_{\lambda_{l-1}}\}$ $\mathfrak{B}$-abhängig ist.

Offenbar ist ein $\mathfrak{B}$-abhängiges System $\mathfrak{S}$ stets auch schwach $\mathfrak{B}$-abhängig. Weiter gelten die folgenden einfachen Aussagen:

(a) *Sei* $r(f_\lambda) = \dim Y_\lambda$ $(\lambda = 1, \cdots, l)$. *Ist dann das System der* f_λ *unabhängig, so ist es auch $\mathfrak{B}$-unabhängig.*

Denn $\{f_1, \cdots, f_l\}(X)$ enthält in diesem Falle eine lokal-analytische Menge der Dimension $\sum_{\lambda=1}^{l} r(f_\lambda)$: Bezeichnet U die Menge der inbezug auf $\{f_1, \cdots, f_l\}$ nichtsingulären Punkte von X, so sei $x_0 \in U$ ein Punkt, in dem X irreduzibel ist und in dem der lokale Rang der holomorphen Abbildung

$\{f_1,\cdots,f_l\}\mid U$ mit ihrem globalen Rang $\sum_{\lambda=1}^{l} r(f_\lambda)$ übereinstimmt; x_0 besitzt eine Umgebung $V(X_0)$, derart, dass $\{f_1,\cdots,f_l\}(V(x_0))$ eine lokal-analytische Menge der Dimension $\sum_{\lambda=1}^{l} r(f_\lambda)$ ist (vgl. [22], Satz 19). Die Bildmenge $\{f_1,\cdots,f_l\}(X)$ kann demnach in keiner analytischen Menge positiver Codimension enthalten sein.

(b) *Ist* f_0 *von* f *𝔅-abhängig, so ist* f_0 *von* f *abhängig.*

Würde nämlich $r(\{f,f_0\})>r(f)$ gelten, so müsste $\{f,f_0\}(X)$ eine lokal-analytische Menge von einer $r(f)$ übertreffenden Dimension enthalten (vgl. die Begründung zu (a)); $\{f,f_0\}(X)$ könnte daher nicht in einer $r(f)$-dimensionalen analytischen Menge enthalten sein, d.h. f_0 wäre von f nicht 𝔅-abhängig.

(b′) *Ist das System* $\mathfrak{S}$ *der* $f_1,\cdots,f_l$ *𝔅-abhängig, so ist es auch abhängig.*

Dies ergibt sich sofort mittels (b).

In besonderen Fällen lässt sich umgekehrt von der Abhängigkeit auf 𝔅-Abhängigkeit schliessen. Die folgenden Propositionen betreffen solche Fälle.

PROPOSITION 6.1.2. *Die meromorphe Abbildung* $f_0\colon X\underset{m}{\longrightarrow}Y_0$ *sei von der meromorphen Abbildung* $f\colon X\underset{m}{\longrightarrow}Y$ *abhängig. Es existiere eine irreduzible analytische Menge* A *in* X*, derart, dass die holomorphe Korrespondenz* $a:=f\mid A\colon A\underset{k}{\longrightarrow}Y$ *eine irreduzible Komponente* $a'\colon A\underset{k}{\longrightarrow}Y$ *besitzt, die eigentlich ist und für die* $r(a')=r(f)$ *gilt. Dann ist* f_0 *von* f *auch 𝔅-abhängig.*

Beweis. Es sei $\mathfrak{F}$ die Klasse aller von f abhängigen meromorphen abbildungen und $f^*\colon X\underset{m}{\longrightarrow}Y^*$ eine $\mathfrak{F}$-maximale meromorphe Abbildung gemäss Satz III. Man hat meromorphe Abbildungen $\beta\colon Y^*\underset{m}{\longrightarrow}Y$ und $\beta_0\colon Y^*\underset{m}{\longrightarrow}Y_0$ mit $f=\beta\bigtriangleup f^*$ und $f_0=\beta_0\bigtriangleup f^*$; nach Zusatz (2) zu Satz III ist β eigentlich und vom Range $r(\beta)=r(f)$. Es ist $\{f,f_0\}=\{\beta\bigtriangleup f^*,\beta_0\bigtriangleup f^*\}=\{\beta,\beta_0\}\bigtriangleup f^*$, wie leicht mittels Proposition 3.4.2 folgt. Mit β ist $\{\beta,\beta_0\}$ nach Proposition 4.3.1 eigentlich, ferner ist $r(\{\beta,\beta_0\})\geqq r(\beta)=\dim Y^*\geqq r(\{\beta,\beta_0\})$, also $r(\{\beta,\beta_0\})=r(\beta)=r(f)$. Demnach ist $\{\beta,\beta_0\}(Y^*)$ eine $r(f)$-dimensionale analytische Menge in $Y\times Y_0$, und $\{f,f_0\}(X)$ ist in $\{\beta,\beta_0\}(Y^*)$ enthalten; f_0 ist daher von f 𝔅-abhängig.

Ist der komplexe Raum X kompakt, so ist für $A:=X$ die Voraussetzung über f in Proposition 6.1.2 stets trivialerweise erfüllt. Damit ergibt sich unmittelbar auch

PROPOSITION 6.1.2a. *Das System* $\mathfrak{S}$ *der meromorphen Abbildungen* $f_\lambda : X \longrightarrow Y_\lambda$ $(\lambda = 1, \cdots, l)$ *sei abhängig; der komplexe Raum* X *sei kompakt. Dann ist* $\mathfrak{S}$ $\mathfrak{B}$*-abhängig.*

Weiter folgt aus der Spezialisierung von Proposition 6.1.2 auf den Fall, dass die analytische Menge A einpunktig ist, sofort

PROPOSITION 6.1.2b. *Das System* $\mathfrak{S}$ *der meromorphen Abbildungen* $f_\lambda : X \xrightarrow[m]{} Y_\lambda$ $(\lambda = 1, \cdots, l)$ *sei im Punkte* $x_0 \in X$ *stark unabhängig. Ist dann* $f_{l+1} : X \xrightarrow[m]{} Y_{l+1}$ *eine meromorphe Abbildung, die von der Abbildung* $\{f_1, \cdots, f_l\}$ *abhängig ist, so ist das System* $\mathfrak{S}_1$ *der Abbildungen* $f_1, \cdots, f_l, f_{l+1}$ $\mathfrak{B}$*-abhängig.*

6.2. Es sei $f : X \xrightarrow[m]{} Y$ eine meromorphe Abbildung, $\mathfrak{F}$ eine f-Klasse von meromorphen Abbildungen. Enthält $\mathfrak{F}$ die identische Abbildung von X, so nennen wir $\mathfrak{F}$ eine *I-Klasse auf X*. $\mathfrak{F}$ heisse ferner *saturiert*, wenn gilt: Ist $\phi : X \xrightarrow[m]{} {}'Y$ eine Abbildung aus $\mathfrak{F}$ und $\psi : {}'Y \xrightarrow[m]{} Z$ eine meromorphe Abbildung, derart, dass $\psi \triangle \phi : X \xrightarrow[m]{} Z$ definiert ist, so gehört $\psi \triangle \phi$ stets zu $\mathfrak{F}$. Offenbar ist $\mathfrak{F}$ genau dann zugleich I-Klasse und saturiert, wenn $\mathfrak{F}$ jede meromorphe Abbildung von X enthält.

Existiert eine $\mathfrak{F}$-maximale Abbildung $f^* : X \xrightarrow[m]{} Y^*$, so gibt es zu $\phi : X \xrightarrow[m]{} {}'Y$ aus $\mathfrak{F}$ stets eine meromorphe Abbildung ${}'\phi : Y^* \xrightarrow[m]{} {}'Y$, sodass $\phi = {}'\phi \triangle f^*$ gilt. Die so bestimmten Abbildungen ${}'\phi$ bilden, wie unmittelbar ersichtlich, eine I-Klasse ${}'\mathfrak{F}$ auf Y^*, welche vermöge der Zuordnung ${}'\phi \rightarrow {}'\phi \triangle f^* = \phi$ bijektiv auf $\mathfrak{F}$ bezogen ist. ${}'\mathfrak{F}$ ist genau dann saturiert, besteht also aus allen meromorphen Abbildungen von Y^*, wenn $\mathfrak{F}$ saturiert ist.

Wir betrachten nun den Fall, dass f wie in Proposition 6.1.2 die Voraussetzung von Satz III erfüllt, dass also für eine irreduzible analytische Menge $A \subset X$ die holomorphe Korrespondenz $a := f \mid A : A \xrightarrow[k]{} Y$ eine irreduzible eigentliche Komponente $a' : A \xrightarrow[k]{} Y$ vom Range $r(a') = r(f)$ besitzt. Nach Satz III und seinem Zusatz (2) hat man dann eine $\mathfrak{F}$-maximale meromorphe Abbildung $f^* : X \xrightarrow[m]{} Y^*$, derart, dass die meromorphe Abbildung $\beta : Y^* \xrightarrow[m]{} Y$, für welche $f = \beta \triangle f^*$ gilt, eigentlich und vom Range $r(\beta) = r(f) = \dim Y^*$ ist. Es ist $a'(A) = f(X) =: Y'$ eine irreduzible $r(f)$-dimensionale analytische Menge in Y. Gleiches trifft für die Bildmenge $\beta(Y^*)$ zu; da $f(X) = (\beta \triangle f^*)(X) \subset \beta(f^*(X)) = \beta(Y^*)$, ist $f(X) = \beta(Y^*)$. Bezeichnet $\beta' : Y^* \xrightarrow[m]{} Y'$ die Beschränkung von β auf Y', so ist β' eigentlich und vom Range $r(\beta')=\dim Y^*=\dim Y'$; demnach ist (Y^*,β',Y') eine meromorphe Überlagerung von Y'. Damit ist beweisen

PROPOSITION 6.2.1. *Sei* $f: X \xrightarrow[m]{} Y$ *eine meromorphe Abbildung und* $\mathfrak{F}$ *eine f-Klasse meromorpher Abbildungen. Es existiere eine irreduzible analytische Menge* A *in* X, *derart, dass die holomorphe Korrespondenz* $a:=f\,|\,A: A \xrightarrow[k]{} Y$ *eine irreduzible Komponente* $a': A \xrightarrow[k]{} Y$ *besitzt, die eigentlich und vom Range* $r(a')=r(f)$ *ist. Dann ist* $f(X)=:Y'$ *eine irreduzible* $r(f)$*-dimensionale analytische Menge in* Y, *und es gibt eine meromorphe Überlagerung* (Y^*,β',Y') *von* Y', *eine surjektive meromorphe Abbildung* $f^*: X \xrightarrow[m]{} Y^*$ *sowie eine I-Klasse* $'\mathfrak{F}$ *meromorpher Abbildungen von* Y^*, *sodass folgendes gilt: Durch die Zuordnung* $'\phi \to {'\phi} \triangle f^*$ $({'\phi} \in {'\mathfrak{F}})$ *wird eine bijektive Beziehung zwischen* $'\mathfrak{F}$ *und* $\mathfrak{F}$ *gegeben;* $'\mathfrak{F}$ *besteht genau dann aus allen meromorphen Abbildungen von* Y^*, *wenn* $\mathfrak{F}$ *saturiert ist. Die Abbildung* f^* *ist* $\mathfrak{F}$*-maximal.*

6.3. Wir interpretieren in diesem Abschnitt die Aussagen über Abhängigkeit meromorpher Abbildungen für den Spezialfall meromorpher Funktionen.

Sei X ein irreduzibler komplexer Raum. Unter einer meromorphen Funktion auf X wird eine meromorphe Abbildung $\phi: X \xrightarrow[m]{} \boldsymbol{P}^1$ in die Riemannsche Zahlenkugel $\boldsymbol{P}^1$ verstanden, die X nicht konstant in den Punkt ∞ von $\boldsymbol{P}^1$ überführt. (Zur Bezeichnung: $\boldsymbol{P}^1$ wird identifiziert mit dem komplex-projektiven Raum der homogenen komplexen Zahlenpaare (u_0,u_1); es ist $\infty=\{(0,u_1)\}\,(u_1\neq 0)$, $\boldsymbol{P}^1-\{\infty\}=\boldsymbol{C}$, $\boldsymbol{P}^{(l)}:=\underbrace{\boldsymbol{P}^1\times\cdots\times\boldsymbol{P}^1}_{l}$).—Die Verknüpfungen meromorpher Funktionen lassen sich wie folgt erklären: Es gibt meromorphe Abbildungen v_i $(i=1,\cdots,4)$ von $\boldsymbol{P}^{(2)}$ in $\boldsymbol{P}^1$, durch die die Addition bezw. Subtraktion bezw. Multiplikation bezw. Division komplexer Zahlen beschrieben wird (ist $(x,y)\in\boldsymbol{C}^2\subset\boldsymbol{P}^{(2)}$, so ist z. B. $v_1(x,y)=x+y\in\boldsymbol{C}\subset\boldsymbol{P}^1$). Sind nun ϕ_1, ϕ_2 meromorphe Funktionen auf X, so sind Summe, Differenz, Produkt und, falls ϕ_2 nicht konstant gleich Null ist, Quotient von ϕ_1 und ϕ_2 jeweils durch $v_i \triangle \{\phi_1,\phi_2\}$ mit dem entsprechenden v_i definiert.

Es gilt:

(1) *Sei* $f: X \xrightarrow[m]{} Y$ *eine meromorphe Abbildung. Dann bildet die Menge der von* f *abhängigen meromorphen Funktionen auf* X *einen Körper* $\mathfrak{M}_f(X)$.

Ist speziell $\iota\colon X \to X$ die identische Abbildung, so schreiben wir $\mathfrak{M}_\iota(X) =: \mathfrak{M}(X)$.

(1′) *Sei $A \subset X$ eine irreduzible analytische Menge, die ein Fundamentalsystem von offenen irreduziblen Umgebungen in X besitzt. Sei $f\colon X \xrightarrow[m]{} Y$ eine meromorphe Abbildung und $\tilde{f}$ der durch f repräsentierte meromorphe Abbildungskeim auf A. Dann bildet die Menge der von $\tilde{f}$ abhängigen meromorphen Funktionskeime auf A einen Körper $\mathfrak{M}_{\tilde{f}}(A, X)$.*

Ein System von l meromorphen Funktionen $\phi_\lambda\colon X \xrightarrow[m]{} \boldsymbol{P}^1$ $(\lambda = 1, \cdots, l)$ heisst algebraisch abhängig, wenn es eine durch ein nichtkonstantes Polynom bestimmte meromorphe Abbildung $\Phi\colon \boldsymbol{P}^{(l)} \xrightarrow[m]{} \boldsymbol{P}^1$ gibt, sodass

$$\Phi \triangle \{\phi_1, \cdots, \phi_l\}\colon X \xrightarrow[m]{} \boldsymbol{P}^1$$

die konstante Abbildung ist, die X in den Nullpunkt des $\boldsymbol{P}^1$ überführt. (Es ist klar, dass jedes Polynom $F(z_1, \cdots, z_l)$ eine meromorphe Abbildung Φ von $\boldsymbol{P}^{(l)}$ in $\boldsymbol{P}^1$ bestimmt; für jeden Punkt $x \in X$, der nicht Unbestimmtheitsstelle oder Pol eines ϕ_λ ist, ist $\Phi \triangle \{\phi_1, \cdots, \phi_l\}(x) = F(\phi_1(x), \cdots, \phi_l(x))$).

Wir stellen nun fest

(2) *Für ein System meromorpher Funktionen sind die Begriffe "schwach $\mathfrak{B}$-abhängig" und "algebraisch abhängig" äquivalent.*

Denn ist das System $\mathfrak{S}$ der meromorphen Funktionen $\phi_\lambda\colon X \xrightarrow[m]{} \boldsymbol{P}^1$ $(\lambda = 1, \cdots, l)$ algebraisch abhängig, ist also $\Phi \triangle \{\phi_1, \cdots, \phi_l\}(X) = \{0\} \in P^1$ (Φ hat die oben angegebene Bedeutung), so muss $\{\phi_1, \cdots, \phi_l\}(X)$ in der 1-codimensionalen analytischen Menge $\Phi^{-1}(0)$ enthalten sein; $\mathfrak{S}$ ist mithin schwach $\mathfrak{B}$-abhängig. Ist umgekehrt $\mathfrak{S}$ schwach $\mathfrak{B}$-abhängig, so heisst dies, dass $\{\phi_1, \cdots, \phi_l\}(X)$ in einer mindestens 1-codimensionalen analytischen Menge A in $\boldsymbol{P}^{(l)}$ enthalten ist. Nach W. L. Chow [9] ist eine solche Menge A stets algebraisch, also gibt es eine durch ein nichtkonstantes Polynom $F(z_1, \cdots, z_l)$ bestimmte meromorphe Abbildung $\Phi\colon \boldsymbol{P}^{(l)} \xrightarrow[m]{} \boldsymbol{P}^1$, sodass $\Phi \triangle \{\phi_1, \cdots, \phi_l\}(X) = \{0\} \in \boldsymbol{P}^1$ gilt.

Aus den Proposition 6.1.2a und 6.1.2b ergeben sich daher die folgenden Sätze von W. Thimm [34, 35, 36, 37]:

(a) *Auf einem kompakten komplexen Raum ist ein System abhängiger meromorpher Funktionen stets algebraisch abhängig.*

(b) *Das System $\mathfrak{S}$ der auf dem komplexen Raum X meromorphen Funktionen $\phi_1, \cdots, \phi_l$ sei im Punkte $x_0 \in X$ stark unabhängig. Ist dann ϕ_{l+1}* eine auf *X meromorphe Funktion, derart, dass das System $\mathfrak{S}_1$ der Funktionen $\phi_1, \cdots, \phi_l, \phi_{l+1}$ abhängig ist, so ist $\mathfrak{S}_1$ algebraisch abhängig.*

Damit (b) aus Proposition 6.1.2b gefolgert werden kann, muss nur gezeigt werden, dass aus der Voraussetzung "$\mathfrak{S}_1$ ist ein abhängiges System" hier "ϕ_{l+1} ist von $\{\phi_1,\cdots,\phi_l\}$ abhängig" folgt. Für konstantes ϕ_{l+1} ist dies klar. Ist ϕ_{l+1} nicht konstant, also $r(\phi_{l+1})=1$, so hat man

$$\begin{aligned} r(\{\phi_1,\cdots,\phi_l\}) &\leqq r(\{\phi_1,\cdots,\phi_{l+1}\}) \\ &= r(\{\{\phi_1,\cdots,\phi_l\},\phi_{l+1}\}) \leqq r(\{\phi_1,\cdots,\phi_l\}) + r(\phi_{l+1}) \\ &= r(\phi_1)+\cdots+r(\phi_l)+1 = l+1. \end{aligned}$$

Würde $r(\{\phi_1,\cdots,\phi_{l+1}\})=l+1$ gelten, so wäre $\mathfrak{S}_1$ ein unabhängiges System; demnach ist $r(\{\{\phi_1,\cdots,\phi_l\},\phi_{l+1}\})=r(\{\phi_1,\cdots,\phi_l\})$ und folglich ϕ_{l+1} von $\{\phi_1,\cdots,\phi_l\}$ abhängig.

Sei nun $f\colon X \xrightarrow[m]{} Y$ eine meromorphe Abbildung, zu der eine komplexe m-Basis (f^*, Y^*) existiert. Es bezeichne $\mathfrak{F}$ die Klasse aller von f abhängigen meromorphen Abbildungen, ferner $'\mathfrak{F}$ die Klasse aller meromorphen Abbildungen von Y^*. Nach 6.2 wird durch die Zuordnung $'\phi \to {'\phi} \triangle f^*$ ($'\phi \in {'\mathfrak{F}}$) eine bijektive Beziehung zwischen $'\mathfrak{F}$ und $\mathfrak{F}$ hergestellt. Es ist klar, dass hierdurch insbesondere ein Isomorphismus des Körpers $\mathfrak{M}(Y^*)$ auf den Körper $\mathfrak{M}_f(X)$ induziert wird.

Es werde jetzt über f weiter vorausgesetzt, dass die Bildmenge $f(X) =: Y'$ eine (irreduzible) analytische Menge in Y ist, ferner, dass die zur Beschränkung $f'\colon X \xrightarrow[m]{} Y'$ von f auf Y' gehörige meromorphe Abbildung $\beta'\colon Y^* \xrightarrow[m]{} Y'$, für welche $f' = \beta' \triangle f^*$ gilt, eine meromorphe Überlagerungsabbildung ist (dies tritt insbesondere ein, wenn f wie in Proposition 6.2.1 die Voraussetzung von Satz III erfüllt). Es gilt dann

(3) *Der Funktionenkörper $\mathfrak{M}(Y^*)$—also auch der Körper $\mathfrak{M}_f(X)$—ist zu einer endlichen algebraischen Erweiterung des Körpers $\mathfrak{M}(Y')$ isomorph.*

Dies kann auf folgende Art nachgewiesen werden: Durch $\phi' \to \phi' \triangle \beta'$ ($\phi' \in \mathfrak{M}(Y')$) wird ein Monomorphismus $^*\beta'\colon \mathfrak{M}(Y') \to \mathfrak{M}(Y^*)$ gegeben; $\mathfrak{M}(Y')$ ist daher zu dem Unterkörper $^*\beta'(\mathfrak{M}(Y'))$ von $\mathfrak{M}(Y^*)$ isomorph. Jede Funktion $\psi^* \in \mathfrak{M}(Y^*)$ genügt einer Gleichung

$$\psi^{*b} + \phi_1{}^* \cdot \psi^{*b-1} + \cdots + \phi_b{}^* = 0,$$

wo b die Blätterzahl der meromorphen Überlagerung (Y^*, β', Y') bezeichnet und $\phi_1{}^*,\cdots,\phi_b{}^*$ Funktionen aus $^*\beta'(\mathfrak{M}(Y'))$ sind ($\phi_1{}^*,\cdots,\phi_b{}^*$ sind die elementarsymmetrischen Funktionen der "Zweige" von ψ^* über Y'); also ist $\mathfrak{M}(Y^*)$ endlich algebraisch über $^*\beta'(\mathfrak{M}(Y'))$.

Ein wichtiger Spezialfall liegt vor, wenn $Y' = Y = \boldsymbol{P}^{(l)}$ ist. Dann ist $\mathfrak{M}(Y') = \mathfrak{M}(\boldsymbol{P}^{(l)})$ isomorph zum Körper $\boldsymbol{C}(z_1,\cdots,z_l)$ der rationalen Funktionen von l komplexen Veränderlichen $z_1, \cdots, z_l$; also sind $\mathfrak{M}(Y^*)$ und $\mathfrak{M}_f(X)$ zu einer endlichen algebraischen Erweiterung von $\boldsymbol{C}(z_1, \cdots, z_l)$ isomorph. Dieser Fall tritt z. B. ein, wenn X kompakt ist und $f: X \underset{m}{\longrightarrow} \boldsymbol{P}^{(l)}$ gegeben wird durch l meromorphe Funktionen $\phi_1, \cdots, \phi_l$ die ein maximales unabhängiges System meromorpher Funktionen auf X bilden (d. h. es ist $f = \{\phi_1, \cdots, \phi_l\}$, wobei das System der $\phi_1, \cdots, \phi_l$ unabhängig, für jede weitere meromorphe Funktion ϕ auf X das System der $\phi_1, \cdots, \phi_l, \phi$ jedoch abhängig ist). Damit haben wir den

Satz von Chow-Thimm.[12] *Der Körper der auf einem irreduziblen kompakten komplexen Raum meromorphen Funktionen ist isomorph zu einer endlichen algebraischen Erweiterung eines rationalen Funktionenkörpers über* $\boldsymbol{C}$.

Wird nur vorausgesetzt, dass Y^* kompakt ist, so ergibt sich weiter, dass $\mathfrak{M}_f(X)$ auch in diesem Falle zu einer endlichen algebraischen Erweiterung eines rationalen Funktionenkörpers über $\boldsymbol{C}$ isomorph ist. Der komplexe Raum Y^* ist z. B. dann kompakt, wenn f eine Unbestimmtheitsstelle maximaler Stufe besitzt. Folglich gilt:

Proposition 6.3.1. *Sei* $f: X \underset{m}{\longrightarrow} Y$ *eine meromorphe Abbildung mit einer Unbestimmtheitsstelle* $x_0 \in X$ *maximaler Stufe. Dann ist der Körper der von* f *abhängigen meromorphen Funktionen auf* X *isomorph zu einer endlichen algebraischen Erweiterung eines rationalen Funktionenkörpers über* $\boldsymbol{C}$.

Für den Fall, dass $Y = \boldsymbol{P}^{(l)}$ und $r(f) = l$ ist (d. h. wenn f durch ein in x_0 stark unabhängiges System von l meromorphen Funktionen gegeben wird), ergibt sich ein bekannter Satz von W. Thimm [34, 36]. Dieser Satz und der Satz von Chow-Thimm wurden auch von R. Remmert [24] durch Spezialisierung einer allgemeineren Aussage, über abhängige meromorphe Funktionen, gewonnen.

Eine der Proposition 6.3.1 entsprechende Aussage, die analog abzuleiten ist, gilt auch für meromorphe Funktionskeime (vgl. auch hierzu [24]):

Proposition 6.3.1′. *Sei* $f: X \underset{m}{\longrightarrow} Y$ *eine meromorphe Abbildung mit einer Unbestimmtheitsstelle* $x_0 \in X$ *maximaler Stufe;* x_0 *besitze ein Fundamentalsystem irreduzibler offener Umgebungen in* X. *Es sei* $\bar{f}$ *der durch* f *repräsentierte meromorphe Abbildungskeim in* x_0. *Dann ist der Körper der von* $\bar{f}$ *abhängigen meromorphen Funktioneskeime in* x_0 *isomorph zu einer*

---

[12] Vgl. [9a, 37]. Ein einfacher Beweis des Satzes wurde von R. Remmert in [21] gegeben; dort finden sich auch Hinweise auf weitere Literatur.

endlichen algebraischen Erweiterung eines rationalen Funktionenkörpers über ***C.***

MÜNCHEN.

LITERATUR.

[1] A. Andreotti and W. Stoll, "Extension of holomorphic maps," *Annals of Mathematics*, vol. 72 (1960), pp. 312-349.

[2] H. Behnke, "Funktionentheorie auf komplexen Mannigfaltigkeiten," *Proceedings of the International Congress of Mathematicians*, 1954, vol. III, pp. 45-57.

[3] ——— and H. Grauert, "Analysis in non-compact complex spaces," in *Analytic Functions*, Princeton, 1960.

[4] A. Borel and A. Haefliger, "La classe d'homologie fondamentale d'un espace analytique," *Bulletin de la Société Mathématique de France*, vol. 89 (1961), pp. 461-513.

[5] N. Bourbaki, Théorie des ensembles, Chap. I-II, 2. Éd., Paris, 1960.

[5a] ———, Topologie générale, Chap. I-II, 3. Éd., Paris, 1961.

[6] H. Cartan, *Séminaire É. N. S.*, 1951/52.

[7] ———, *Séminaire É. N. S.*, 1953/54.

[8] ———, "Quotients of complex analytic spaces," *International Colloquium on Function Theory*, Tata Institute Bombay, January 1960, pp. 1-15.

[9] W. L. Chow, "On compact analytic varieties," *American Journal of Mathematics*, vol. 71 (1949), pp. 893-914.

[9a] ——— and K. Kodaira, "On analytic surfaces with two independent meromorphic functions," *Proceedings of the National Academy of Sciences*, vol. 38 (1952), pp. 319-325.

[10] H. Grauert, "Ein Theorem der analytischen Garbentheorie und die Modulräume komplexer Strukturen," *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, No. 5 (1960), pp. 233-292.

[10a] ———, "Über Modifikationen und exceptionelle analytische Mengen," *Mathematische Annalen*, vol. 146 (1962), pp. 331-368.

[11] ——— and R. Remmert, "Zur Theorie der Modifikationen. I. Stetige und eigentliche Modifikationen komplexer Räume," *Mathematische Annalen*, vol. 129 (1955), pp. 274-296.

[12] ———, "Komplexe Räume," *Mathematische Annalen*, vol. 136 (1958), pp. 245-318.

[13] K. H. Hedtfeld, "Starre einfach-zusammenhängende Holomorphiegebiete," *Schriftenreihe des Mathematischen Instituts der Universität Münster*, Heft 8 (1954).

[14] H. Kerner, "Holomorphiehüllen zu K-vollständigen komplexen Räumen," *Mathematische Annalen*, vol. 138 (1960), pp. 316-328.

[15] K. Koch, "Die analytische Projektion," *Schriftenreihe des Mathematischen Instituts der Universität Münster*, Heft 6 (1953).

[16] N. Kuhlmann, "Zur Theorie der Modifikationen algebraischer Varietäten," *Schriftenreihe des Mathematischen Instituts der Universität Münster*, Heft 14 (1959).

[17] ———, "Projektive Modifikationen komplexer Räume," *Mathematische Annalen*, vol. 139 (1960), pp. 217-238.

[18] ———, "Die Normalisierung komplexer Räume," *Mathematische Annalen*, vol. 144 (1961), pp. 110-125.

[18a] ———, "Über die Auflösung der Singularitäten 3-dimensionaler komplexer Räume I," *Mathematische Annalen*, vol. 151 (1963), pp. 304-331.

[18b] R. Narasimhan, "Imbedding of holomorphically complete complex spaces," *American Journal of Mathematics*, vol. 82 (1960), pp. 917-934.

[19] K. Oka, "Sur les fonctions analytiques de plusieurs variables. VIII. Lemme fondamental," *Journal of the Mathematical Society of Japan*, vol. 3 (1951), pp. 204-214 and 259-278.

[20] R. Remmert, "Projektionen analytischer Mengen," *Mathematische Annalen*, vol. 130 (1956), pp. 410-441.

[20a] ———, "Über stetige und eigentliche Modifikationen komplexer Räume," *Colloque de Topologie de Strasbourg*, Déc. 1954, pp. 1-17.

[21] ———, "Meromorphe Funktionen in kompakten komplexen Räumen," *Mathematische Annalen*, vol. 132 (1956), pp. 277-288.

[22] ———, "Holomorphe und meromorphe Abbildungen komplexer Räume," *Mathematische Annalen*, vol. 133 (1957), pp. 328-370.

[23] ———, "Sur les espaces analytiques holomorphiquement séparables et holomorphiquement convexes," *Comptes Rendus*, vol. 243 (1956), pp. 118-121.

[24] ———, "Analytic and algebraic dependence of meromorphic functions," *American Journal of Mathematics*, vol. 82 (1960), pp. 891-899

[25] ——— and K. Stein, "Über die wesentlichen Singularitäten analytischer Mengen," *Mathematische Annalen*, vol. 126 (1953), pp. 263-306.

[26] ———, "Eigentliche holomorphe Abbildungen," *Mathematische Zeitschrift*, vol. 73 (1960), pp. 159-189.

[27] J. P. Serre, "Géométrie algébrique et géométrie analytique," *Annales de l'Institut Fourier*, G (1955/56), pp. 1-42.

[28] K. Stein, "Analytische Projektion komplexer Mannigfaltigkeiten," *Colloque sur les fonctions de plusieurs variables*, Bruxelles, 1953, pp. 97-107.

[29] ———, "Analytische Zerlegungen komplexer Räume," *Mathematische Annalen*, vol. 132 (1956), pp. 63-93.

[30] ———, "Die Existenz komplexer Basen zu holomorphen Abbildungen," *Mathematische Annalen*, vol. 136 (1958), pp. 1-8.

[31] W. Stoll, "Über meromorphe Modifikationen I-III," *Mathematische Zeitschrift*, vol. 61 (1954), pp. 204-234, 467-488 and vol. 62 (1955), pp. 189-210.

[32] ———, "Über meromorphe Modifikationen IV, V," *Mathematische Annalen*, vol. 130 (1955), pp. 147-182, 272-316.

[33] ———, "Über meromorphe Abbildungen komplexer Räume I, II," *Mathematische Annalen*, vol. 136 (1958), pp. 201-239, 393-429.

[34] W. Thimm, "Über ausgeartete meromorphe Abbildungen II," *Mathematische Annalen*, vol. 125 (1952), pp. 264-283.

[35] ———, "Über die Menge der singulären Bildpunkte einer meromorphen Abbildung," *Mathematische Zeitschrift*, vol. 57 (1953), pp. 456-480.

[36] ———, "Untersuchungen über ausgeartete meromorphe Abbildungen," *Mathematische Annalen*, vol. 127 (1954), pp. 150-161.

[37] ———, "Meromorphe Abbildungen von Riemannschen Bereichen," *Mathematische Zeitschrift*, vol. 60 (1954), pp. 435-457.

[38] K. Wolffhardt, "Existenzbedingungen für maximale holomorphe und meromorphe Abbildungen," *Dissertation*, München, 1963.

ON THE ALGEBRA OF REPRESENTATIVE FUNCTIONS OF AN ANALYTIC GROUP, II.

By G. Hochschild and G. D. Mostow.*

1. Introduction. We continue our study [1] of the algebra $R(G)$ of the representative functions of an analytic group G. In particular, we employ $R(G)$ for investigating the algebraic hulls of G. For the usual algebraic-geometric reasons, we confine our attention to *complex* analytic groups.

The algebraic hulls of G are viewed as the groups of all proper automorphisms of finitely generated fully stable subalgebras of $R(G)$ separating the points of G. One of our results (Th. 3.1) is that the structure of an algebraic hull and the location of G in it are particularly transparent when the corresponding subalgebra of $R(G)$ contains a normal basic subalgebra. Conversely, certain comparatively mild regularity requirements on an algebraic hull imply that the corresponding subalgebra of $R(G)$ contains a normal basic subalgebra (Th. 4.1) or, equivalently, that the algebraic hull "dominates," by a rational epimorphism, a "nuclear" algebraic hull $A(G,K)$ associated in a natural way with every nucleus K of G.

The nuclear algebraic hulls are shown to be the solution of a certain universal problem concerning analytic homomorphisms of G into algebraic linear groups whose restrictions to the given nucleus K satisfy a compatibility condition with respect to the adjoint representation of K (Th. 5.3). The structure of the nuclear algebraic hulls $A(G,K)$ and of the corresponding subalgebras of $R(G)$ is analyzed in detail.

2. Auxiliary representation theory. We assemble a few results of representation theory for later use. The first of these is a generalization of the usual weight space decomposition of a representation space for a nilpotent Lie algebra.

Let L be a finite-dimensional solvable Lie algebra over an algebraically closed field F of characteristic 0. Let ρ be a representation of L by linear endomorphismsof a finite-dimensional F-space V, and let ρ' denote the semisimple representation of L that is associated with ρ. Then ρ' is a direct sum of 1-dimensional representations, which may be regarded as Lie algebra

---

Received April 12, 1964.

* Performed under NSF-GP-1662.

homomorphism $L \rightarrow F$. We denote by $P(\rho)$ the finite set of these Lie algebra homomorphisms. Let α stand for the adjoint representation of L. We partition the set $P(\rho)$ into equivalence classes, calling two elements h_1 and h_2 of $P(\rho)$ equivalent if $h_1 - h_2$ belongs to the additive group of Lie algebra homomorphisms $L \rightarrow F$ that is generated by $P(\alpha)$.

If h is any Lie algebra homomorphism $L \rightarrow F$ and x is an element of L we denote by $\rho_h(x)$ the linear endomorphism of V that is defined by $\rho_h(x)(v) = \rho(x)(v) - h(x)v$. If γ is an equivalence class of elements of $P(\rho)$ we denote by V_γ the set of all elements v of V for which there exist elements $h_1, \cdots, h_n$ in γ such that $\rho_{h_1}(x_1) \cdots \rho_{h_n}(x_n)(v) = 0$ for all n-tuples $(x_1, \cdots, x_n)$ of elements of L. One verifies immediately that each V_γ is an L-stable subspace of V.

LEMMA 2.1. *In the notation introduced above, V is the direct sum of the L-stable subspaces V_γ, where γ ranges over the equivalence classes of elements of $P(\rho)$.*

Proof. We may evidently choose an element x in L such that the values $h(x)$, as h ranges over the finite set $((0) \cup P(\rho)) + ((0) \cup P(\alpha))$, are all distinct. If e is any linear endomorphism of V and a is an element of F we denote by $V_a(e)$ the subspace of V consisting of all elements that are annihilated by some power of $e - aI$, where I stands for the identity map on V. We claim that if $V_a(\rho(x))$ is not (0) then $a \in P(\rho)(x)$, i. e., $a = h(x)$ with some h in $P(\rho)$. Indeed, our assumption implies that there is a non-zero element v in V such that $\rho(x)(v) = av$. On the other hand, it is clear from the definition of $P(\rho)$ that there are elements $h_1, \cdots, h_n$ in $P(\rho)$ such that $\rho_{h_1}(x) \cdots \rho_{h_n}(x)(v) = 0$. This gives $(a - h_1(x)) \cdots (a - h_n(x))v = 0$, so that we must have $a = h_i(x)$ for at least one of the indices i.

Since V is evidently the sum of the subspaces $V_a(\rho(x))$, we have therefore $V = \sum_{a \in P(\rho)(x)} V_a(\rho(x))$. Similarly, $L = \sum_{a \in P(\alpha)(x)} L_a(\alpha(x))$. If γ is any equivalence class of elements of $P(\rho)$ we put $W_\gamma = \sum_{a \in \gamma(x)} V_a(\rho(x))$. It is clear from the above that V is the sum of the subspaces W_γ. Moreover, this sum is evidently a direct sum. Hence it will suffice to show that $W_\gamma = V_\gamma$ for each γ.

For this purpose, we show first that W_γ is L-stable. By the above decomposition of L, it will suffice to show that $\rho(y)(W_\gamma) \subset W_\gamma$ if $y \in L_b(\alpha(x))$ with b in $P(\alpha)(x)$. Now we have

$$(\rho(x) - a)\rho(y) = \rho([x, y]) + \rho(y)(\rho(x) - aI),$$

whence

$$(\rho(x)-(a+b))\rho(y)=\rho((\alpha(x)-bI)(y))+\rho(y)(\rho(x)-aI).$$

Hence we obtain, by induction on n,

$$(\rho(x)-(a+b)I)^n\rho(y)=\sum_{k=0}^{n}\binom{n}{k}\rho((\alpha(x)-bI)^k(y))(\rho(x)-aI)^{n-k},$$

which shows that $\rho(y)(V_a(\rho(x)))\subset V_{a+b}(\rho(x))$. Now suppose that $a=h(x)$, with h in γ, and $b=k(x)$, with k in $P(\alpha)$. If $V_{a+b}(\rho(x))\neq(0)$ then we must have $a+b=t(x)$, with t in $P(\rho)$. By the choice of x, it follows that $h+k=t\in P(\rho)$, so that $h+k\in\gamma$ and $a+b\in\gamma(x)$. Thus we conclude from our above result that $\rho(y)(W_\gamma)\subset W_\gamma$, and so we have shown that W_γ is L-stable.

Let ρ^γ denote the representation of L on V_γ that is induced by ρ. Then we have $P(\rho^\gamma)\subset\gamma$ and hence

$$V_\gamma=\sum_{a\in P(\rho^\gamma)(x)}(V_\gamma)_a(\rho^\gamma(x))\subset\sum_{a\in\gamma(x)}V_a(\rho(x))=W_\gamma.$$

Now consider the factor module W_γ/V_γ. If this were not (0) it would contain a 1-dimensional L-stable subspace Fu, say. Then we have, for every element y of L, $y\cdot u=h(y)u$, where h belongs to $P(\rho)$. On the other hand, if v is a representative of u in W_γ, there are elements $h_1,\cdots,h_n$ in γ (not necessarily distinct) such that $\rho_{h_1}(x)\cdots\rho_{h_n}(x)(v)=0$. If we pass over to W_γ/V_γ we obtain from this that $(h(x)-h_1(x))\cdots(h(x)-h_n(x))u=0$, whence $h(x)=h_i(x)$ for some index i. By the choice of x, this implies that $h=h_i\in\gamma$. Since $\rho(y)(v)-h(y)v\in V_\gamma$ for every element y of L, this gives $v\in V_\gamma$, whence the contradiction $u=0$. Thus we must have $V_\gamma=W_\gamma$, and Lemma 2.1 is proved.

COROLLARY 2.1. *Let G be a solvable complex analytic group, and let σ be a complex analytic representation of G by linear automorphisms of a finite-dimensional complex vector space V. Then the analogue of Lemma 2.1 holds for σ.*

Proof. Let $P(\sigma)$ denote the set of complex analytic homomorphisms $G\to C^*$ that is determined by the semisimple representation associated with σ, where C^* denotes the multiplicative group of the non-zero complex numbers. If ϕ is any complex analytic representation of G denote by $\phi^\cdot$ the differential of ϕ, i.e., the induced representation of the Lie algebra $G^\cdot$ of G. Then the map that sends each ϕ onto $\phi^\cdot$ is evidently a bijective map of $P(\sigma)$ onto $P(\sigma^\cdot)$.

Let β be the adjoint representation of G. Then it is clear that two elements of $P(\sigma)$ are equivalent mod. $P(\beta)$ if and only if their differentials are equivalent mod. $P(\beta^{\cdot})$. Our corollary will therefore follow from Lemma 2.1, applied to $\sigma^{\cdot}$, as soon as we have shown that, for each equivalence class δ of elements of $P(\sigma)$, we have $V_\delta = V_{\delta^{\cdot}}$.

Let q be any complex analytic homomorphism $G \to C^*$ and let τ be any finite-dimensional complex analytic representation of G. We claim that the endomorphisms $\tau(u) - q(u)I$ are 0 for all elements u of G if and only if the endomorphisms $\tau(x) - q(x)I$ are 0 for all elements x of $G^{\cdot}$. In fact, the first condition means that the representation τ_q defined by $\tau_q(u) = q(u)^{-1}\tau(u)$ is trivial, which is the case if and only if its differential is trivial. Since $(\tau_q)^{\cdot}(x) = \tau^{\cdot}(x) - q^{\cdot}(x)I$ for every element x of $G^{\cdot}$, our claim is established. This enables us to show, by means of an evident inductive procedure, that we have indeed $V_{\delta^{\cdot}} = V_\delta$.

Let G be a group, F a field. We say that an F-algebra A of F-valued representative functions on G is *fully stable* if it is stable under the left and right translations as well as under the involution $g \to g'$, where $g'(x) = g(x^{-1})$. We shall require the following elementary representation theoretical lemma, which is not new, but is included here for the convenience of the reader.

LEMMA 2.2. *Let F be an arbitrary field, G a group. Let A be a fully stable F-algebra of F-valued representative functions on G, and let V be any non-zero left G-submodule of A. Then the constant function on G with value 1 belongs to the ideal of A that is generated by V.*

Proof. Without loss of generality, we may assume that V is simple as a left G-module. Then the dual G-module V^* is also simple. The stability assumption on A implies that the representative functions associated with V^* belong to A. Now V^* is isomorphic with a G-submodule of the direct sum of a finite number of copies of the G-module constituted by the representative functions associated with V^*. Since V^* is simple, this implies that there is a G-module monomorphism $\phi\colon V^* \to A$.

Now let $(v_1, \cdots, v_n)$ be an F-basis of V such that $v_1(1) = 1$ and $v_i(1) = 0$ for all the other indices i. Let $(\gamma_1, \cdots, \gamma_n)$ be the dual basis of V^*. For x in G, write $x \cdot v_i = \sum_{j=1}^{n} f_{ji}(x) v_j$. Then we have $(x \cdot \gamma_k)(v_i) = \gamma_k(x^{-1} \cdot v_i) = f_{ki}(x^{-1})$, so that $x \cdot \gamma_k = \sum_{j=1}^{n} f_{kj}(x^{-1}) \gamma_j$.

Now put $a_k = \phi(\gamma_k) \in A$. Choose y in G such that $a_1(y) \neq 0$. For every x in G, we have $x \cdot a_k \cdot y = \sum_{j=1}^{n} f_{kj}(x^{-1}) a_j \cdot y$. Hence

$$x \cdot (\sum_{k=1}^{n} (a_k \cdot y) v_k) = \sum_{k,j,t} f_{kj}(x^{-1}) (a_j \cdot y) f_{tk}(x) v_t = \sum_{t=1}^{n} (a_t \cdot y) v_t.$$

Thus the function $\sum_{k=1}^{n} (a_k \cdot y) v_k$ is constant and, by evaluating it at 1, we find that its value is $a_1(y) \neq 0$. Hence the constant function with value 1 is equal to $a_1(y)^{-1} \sum_{k=1}^{n} (a_k \cdot y) v_k$, which evidently belongs to the ideal of A that is generated by V. This completes the proof of Lemma 2.2.

3. Regular subalgebras.

Let G be a complex analytic group, and denote by $R(G)$ the algebra of all analytic representative functions on G. We shall assume throughout that $R(G)$ separates the points of G. A subalgebra of $R(G)$ will be called a *regular subalgebra* if it is fully stable and finitely generated as a C-algebra and contains a normal basic subalgebra of $R(G)$. As usual, we put $Q = \exp(\mathrm{Hom}(G, C))$, where $\mathrm{Hom}(G, C)$ stands for the group of all analytic homorphisms of G into the additive group C of the complex number field.

LEMMA 3.1. *Let S be a regular subalgebra of $R(G)$, and let B be a normal basic subalgebra of $R(G)$ that is contained in S. Then $S = B[S \cap Q]$.*

Proof. Every non-zero element of S can be written in one and only one way as a sum $\sum_{i=1}^{n} b_i q_i$, where the b_i's are non-zero elements of B and the q_i's are distinct elements of Q. Our lemma amounts to the statement that then each q_i belongs to S. Suppose this is false, and consider non-zero elements of S in which none of the q_i's belongs to S. Among these, take one in which n is minimal. Let N denote the radial of the commutator subgroup of G, and recall that every complex analytic representation of G induces a unipotent representation of N and that, for every element x of N and every element q of Q, we have $x \cdot q = q$. By applying a finite sequence of left operations $x - 1$, with x in N, to our element of S in succession, we obtain an element $\sum_{i=1}^{n} c_i q_i$ of S in which each c_i belongs to the N-fixed subalgebra B^N of B and, because of the minimality of n, is different from 0. By [1, Lemma 4.1], we have $B^N = R(G)^K[\mathrm{Hom}(G, C)]$, where K is the nucleus of G that is associated with B. In particular, this shows that B^N is stable under both the left and the right translations.

Now let x be an element of G. Translating our element of S from the left by x and multiplying by c_1, we obtain

$$\sum_{i=1}^{n} (x\cdot c_i)\, q_i(x)\, c_1 q_i \in S.$$

On the other hand, multiplying our element of S by $(x\cdot c_1)\, q_1(x)$, we obtain

$$\sum_{i=1}^{n} (x\cdot c_1)\, q_1(x)\, c_i q_i \in S.$$

Subtracting and using the minimality of n, we find that, for each i,

$$(x\cdot c_1)\, q_1(x)\, c_i = (x\cdot c_i)\, q_i(x)\, c_1.$$

Thus, for every element y of G,

$$c_1(yx)\, q_1(x)\, c_i(y) = c_i(yx)\, q_i(x)\, c_1(y).$$

Hence we have $c_i(y)\,(c_1\cdot y)\, q_1 = c_1(y)\,(c_i\cdot y)\, q_i$. Here, the coefficients of q_1 and q_i are element of B^N. Since the elements of Q are free over B, it follows that we must have $c_i(y)\,(c_1\cdot y) = 0$ whenever $i \neq 1$. Choosing y so that $c_i(y)\, c_1(y) \neq 0$, we obtain a contradiction from this by evaluating at 1, except in the case where $n = 1$. Thus we must have $n = 1$ and $c_1 q_1 \in S$.

Now let V be the left G-submodule of S that is generated by $c_1 q_1$. Then we have $V \subset S q_1$, and it follows from Lemma 2.2 that there is an element s in S such that $s q_1 = 1$. Hence $q_1^{-1} = s \in S$. Since S is stable under the involution $f \rightarrow f'$, this gives $q_1 \in S$, contrary to our assumptions. Thus $S = B[S \cap Q]$.

LEMMA 3.2. *Let S be any fully stable subalgebra of $R(G)$ separating the points of G. Then S contains the representative functions associated with the adjoint representation of G.*

Proof. Without loss of generality, we may assume that S is finitely generated.

Let A_S denote the group of all proper automorphisms of S, and recall the following standard facts. A_S is an algebraic linear group, representable as such on any left translation stable finite-dimensional generating subspace of S. If we associate with an element s of S the function s^* on A_S defined by $s^*(\alpha) = \alpha(s)(1)$ then the map $s \rightarrow s^*$ is an isomorphism of S onto the algebra of all rational representative functions on A_S, because S is fully stable. The group G_S of the left translations of S is a normal analytic subgroup of A_S, and G_S is algebraically dense in A_S. Since S separates the points of G, we may identify G with G_S by the map that associates with each element of G the left translation it effects on S. When this is done, every element s of S becomes identified with the restriction of s^* to G_S. Since G_S is a normal analytic subgroup of A_S, the Lie algebra of G_S is an ideal of the Lie algebra of A_S. Now the adjoint representation of A_S is a rational representation, and the adjoint representation of G is the restriction to G_S of the adjoint representation of A_S on the Lie algebra of G_S. Hence the representative functions associated with the adjoint representation of G are the restrictions to G_S of rational representative functions on A_S and therefore belong to S. This completes the proof of Lemma 3.2.

THEOREM 3.1. *Let S be a regular subalgebra of $R(G)$, and let M be a maximal fully reducible subgroup of the algebraic group A_S of all proper automorphisms of S. Let G_S be the canonical image of G in A_S. Then G_S is an algebraically dense and topologically closed normal analytic subgroup of A_S, and $M \cap G_S$ is an algebraic subgroup of M and a maximal reductive analytic subgroup of G_S. There is an algebraic subgroup X of M such that M is the direct product $X \times (M \cap G_S)$, in the sense of algebraic groups, and A_S is the semidirect product $X \cdot G_S$, in the sense of complex analytic groups. Moreover, the algebraic group X is isomorphic with $(C^*)^d$, where C^* denotes the multiplicative group of the non-zero complex numbers and d is the rank of the finitely generated free abelian group $S \cap Q$.*

Proof. Since S is finitely generated as a C-algebra, we see immediately from Lemma 3.1 that $S \cap Q$ is finitely generated as a group. Since the group Q has no elements of finite order, it follows that $S \cap Q$ is a finitely generated free abelian group. Hence the group $\mathrm{Hom}(S \cap Q, C^*)$ of all homomorphisms of $S \cap Q$ into C^* is isomorphic with $(C^*)^d$ by an isomorphism obtained in the evident fashion from a basis of $S \cap Q$. Accordingly, we regard $\mathrm{Hom}(S \cap Q, C^*)$ as a complex analytic group or also as an algebraic group.

Let $\alpha \in A_S$, $q \in S \cap Q$, $f \in \mathrm{Hom}(G, C)$. Then we have $\alpha(q) = \alpha'(q)q$, with $\alpha'(q) \in C^*$, and $\alpha(f) = f + \alpha'(f)$, with $\alpha'(f) \in C$. Let f_q be the unique element of $\mathrm{Hom}(G, C)$ such that $\exp(f_q) = q$. We define an element α^* of $\mathrm{Hom}(S \cap Q, C^*)$ by setting $\alpha^*(q) = \alpha'(q)\exp(-\alpha'(f_q))$, for every element q of $S \cap Q$. Then one sees immediately that the map $\alpha \rightarrow \alpha^*$ is a complex analytic homomorphism of A_S into $\mathrm{Hom}(S \cap Q, C^*)$, and that G_S lies in the kernel of this homomorphism.

Now let B be a normal basic subalgebra of $R(G)$ that is contained in S, and let B' denote the image of B under the involution $f \rightarrow f'$. Since

$S = B[S \cap Q]$, by Lemma 3.1, we have also $S = B'[S \cap Q]$. On the other hand, $B'[Q] = R(G)$. We recall also that B', like B, contains $\mathrm{Hom}(G, C)$, and that B' is stable under the right translations.

Let α be an element of A_S such that $\alpha^* = 1$. It is clear from what we have just seen that there is a proper automorphism β of $R(G)$ such that β coincides with α on B', while $\beta(q) = \exp(\alpha'(f_q))q$, for every element q of Q. By this very definition, β is a perfect automorphism of $R(G)$ and hence is the left translation effected by an element of G. Since $\alpha^* = 1$ and $S = B'[S \cap Q]$, it is clear that β actually coincides with α on S, so that $\alpha \in G_S$. Thus G_S is precisely the kernel of our homomorphism $\alpha \rightarrow \alpha^*$, and is therefore a closed normal subgroup of A_S.

Now let X be the subgroup of A_S consisting of the automorphisms that leave the elements of B' fixed. Then X is evidently an algebraic subgroup of A_S. We have $X \cap G_S = (1)$, because an automorphism that leaves the elements of B' fixed leaves the elements of $\mathrm{Hom}(G, C)$ fixed so that, if it belongs to G_S, it leaves the elements of $S \cap Q$ fixed.

Now let ϕ be any element of $\mathrm{Hom}(S \cap Q, C^*)$. Then we may define an element α of A_S by demanding that α be the identity on B' and that $\alpha(q) = \phi(q)q$, for every element q of $S \cap Q$. Then $\alpha \in X$ and $\alpha^* = \phi$. Thus our homomorphism $\alpha \rightarrow \alpha^*$ maps X isomorphically onto $\mathrm{Hom}(S \cap Q, C^*)$. Moreover, it is clear from the definition that the restriction of the homomorphism $\alpha \rightarrow \alpha^*$ to X is a rational homomorphism. Thus X is isomorphic with $(C^*)^d$, in the sense of algebraic groups. Since we have evidently $A_S = XG_S$, we have proved that A_S is the semidirect product $X \cdot G_S$, in the sense of complex analytic groups. It is a standard known result that G_S is algebraically dense in A_S.

Let K be the nucleus of G that is associated with the normal basic subalgebra B. By [1, Theorem 4.1], there is an element x in G such that $B \cdot x$ is the normal basic subalgebra that is constructed in the standard fashion [1, pp. 114-115] from a semidirect product decomposition $G = H \cdot K$, where H is a closed reductive analytic subgroup of G. Replacing B by $B \cdot x$, if necessary, we may therefore take B to be such a standard normal basic subalgebra. It is evident from the standard construction [1, pp. 114-115] that then B is stable under the right translations with the elements of H, whence B' is stable under the left translations with the elements of H, i.e., B' is stable with respect to H_S. Moreover, the left translations with the elements of H leave the elements of Q fixed, so that the elements of $S \cap Q$ are H_S-fixed. Hence it is evident that the elements of X commute with the elements of H_S. Since H is reductive, H_S is a fully reducible algebraic subgroup of A_S. Since X and H_S commute element-wise and since $X \cap H_S = (1)$, the group generated by X and H_S is the direct product $X \times H_S$, and it is clearly a fully reducible algebraic subgroup of A_S (noting that a subgroup generated by algebraic subgroups is an algebraic subgroup).

Let Y be a fully reducible subgroup of A_S containing $X \times H_S$. Then $Y = (X \times H_S)(Y \cap K_S)$. Let y be an element of Y and let k be an element of K_S. Then $yky^{-1}k^{-1}$ lies in the commutator subgroup of A_S, which coincides with the commutator subgroup of G_S. The analytic subgroup K_S of A_S is normal in G_S. Since G_S is algebraically dense in A_S, the consideration of the adjoint representation of A_S shows immediately that K_S is normal also in A_S. In particular, $Y \cap K_S$ is a normal subgroup of Y and is therefore fully reducible. Now consider the representation of $Y \cap K_S$ on $C + \mathrm{Hom}(G, C)$. On the one hand, this is evidently a unipotent representation. On the other hand, since $Y \cap K_S$ is fully reducible, this representation is semisimple. Hence $Y \cap K_S$ leaves the elements of $\mathrm{Hom}(G, C)$ fixed. Let N_S denote the radical of the commutator subgroup of G_S. The kernel of the representation of K_S on $C + \mathrm{Hom}(G, C)$ is known to be N_S. We may therefore conclude that $Y \cap K_S \subset N_S$. But N_S is a unipotent subgroup of A_S. Hence $Y \cap K_S = (1)$, and we have shown that $Y = X \times H_S$. Thus $X \times H_S$ is a maximal fully reducible subgroup of A_S. It is known, but follows also from the above argument, that H_S is a maximal reductive analytic subgroup of G_S. Finally, since any two maximal fully reducible subgroups of A_S are conjugate by an inner automorphism, it is clear that the conclusions of Theorem 3.1 hold for every maximal fully reducible subgroup M, provided that they hold for one of them. The proof of Theorem 3.1 is now complete.

4. Regular algebraic hulls. By an *algebraic hull* of the complex analytic group G we shall mean an analytic monomorphism ρ of G into an algebraic linear group L such that $\rho(G)$ is algebraically dense in L. Such an algebraic hull $\rho\colon G \to L$ will be called a *regular algebraic hull* if it satisfies the following two requirements: (1) for every analytic homomorphism ζ of G into C there is a rational homomorphism $\zeta^\rho\colon L \to C$ such that $\zeta^\rho \circ \rho = \zeta$; (2) there is a rational endomorphism ϕ of L such that $\phi \circ \phi = \phi$ and $\phi(L)$ is a maximal reductive analytic subgroup of $\rho(G)$.

It is clear from Theorem 3.1 that if S is any regular subalgebra of $R(G)$ then the canonical map $G \to A_S$ is a regular algebraic hull; the above requirement (1) is satisfied because $\mathrm{Hom}(G, C)$ is contained in S. We wish to prove the converse of this result.

THEOREM 4.1. *Every regular algebraic hull of G can be identified with an algebraic hull $G \to A_S$, where S is a regular subalgebra of $R(G)$.*

Proof. Let $\rho: G \to L$ be a regular algebraic hull of G. If $P(L)$ denotes the algebra of all rational representative functions on L then $P(L) \circ \rho$ is a subalgebra of $R(G)$. Since $\rho(G)$ is algebraically dense in L, the map $f \to f \circ \rho$ of $P(L)$ onto $P(L) \circ \rho$ is an isomorphism. Thus we may identify $P(L)$ with a fully stable subalgebra S of $R(G)$, and we may identify L with the group A_S of all proper automorphism of S, whereby ρ is transported into the canonical map $G \to A_S$. What we have to prove amounts simply to the result that S is a *regular* subalgebra of $R(G)$.

By condition (2), we have $\phi(A_S) = H_S$, where H is a maximal reductive analytic subgroup of G. Now H_S is a fully reducible subgroup of A_S and is therefore contained in a maximal fully reducible subgroup M of A_S. Let K be any nucleus of G. Then G is the semidirect product $H \cdot K$. Hence G_S is the semidirect product $H_S \cdot K_S$, and $M \cap G_S = H_S(M \cap K_S)$. By condition (1), we have $\mathrm{Hom}(G, C) \subset S$. Hence we see exactly as in our proof of Theorem 3.1 (where we showed that $Y \cap K_S = (1)$) that $M \cap K_S = (1)$, so that $M \cap G_S = H_S$.

We know from [1, p. 123] that A_S is the semidirect product $M \cdot U_S$, where U is the unipotent hull of G. Since ϕ is a rational homomorphism, $\phi(U_S)$ is a unipotent normal subgroup of H_S. Since H_S is fully reducible, it follows that $\phi(U_S) = (1)$, whence $H_S = \phi(M)$. Since $\phi \circ \phi = \phi$, it is clear that M is the semidirect product $X \cdot H_S$, where X is the kernel of the restriction of ϕ to M. On the other hand, the commutator subgroup of M lies in $M \cap G_S = H_S$. Hence M is actually the direct product $X \times H_S$, and X is abelian. We have the decomposition $A_S = (X \times H_S) \cdot U_S$. Since $X \cdot U_S$ is contained in the kernel of ϕ, this shows that $X \cdot U_S$ actually coincides with the kernel of ϕ. Since X is abelian, it is clear that $X \cdot U_S$ is solvable.

Now let V denote the kernel of $\phi \circ \rho$. Then $V_S = \rho(V)$ is the kernel of the restriction of ϕ to G_S, i.e., $V_S = (X \cdot U_S) \cap G_S$. Since the restriction of ϕ to G_S is a projection homomorphism of G_S onto H_S, it is clear that G_S is the semidirect product $V_S \cdot H_S$. In particular, V_S is therefore a closed normal complex analytic subgroup of G_S. Since V_S is contained in the solvable complex analytic group $X \cdot U_S$, we conclude that V_S is solvable. Moreover, V_S is homeomorphic with the homogeneous space G_S/H_S and is therefore simply connected. Thus V_S is a nucleus of G_S, and so V is a nucleus of G.

Let W denote the algebraic hull of V_S in A_S. Since $X \cdot U_S$ is an algebraic subgroup of A_S, we have $W \subset X \cdot U_S$. On the other hand, $H_S W$ is an algebraic subgroup of A_S containing $H_S V_S = G_S$, whence $H_S W = A_S$. Thus $X \cdot U_S \subset H_S W$. Since $W \subset X \cdot U_S$ and $H_S \cap (X \cdot U_S) = (1)$, this gives $X \cdot U_S = W$. Hence the V_S-fixed part of S coincides with the $X \cdot U_S$-fixed part of S. Since $X \cdot U_S$ is a normal algebraic subgroup of A_S, the $X \cdot U_S$-fixed part of S is the algebra of all rational representative functions on $A_S/(X \cdot U_S)$ and therefore separates the points of H_S. Regarding S again as an algebra of functions on G, we see that this means that the V-fixed part S^V of S separates the points of G/V. Since G/V is reductive and since S^V is fully stable, this implies that S^V is the algebra of all analytic representative functions on G/V, i. e., $S^V = R(G)^V$. Thus we have $R(G)^V \subset S$.

Now let δ be any proper automorphism of $R(G)$ that leaves the elements of S fixed. By Lemma 3.2, S contains the representative functions associated with the adjoint representation of G. Thus δ leaves the representative functions associated with the adjoint representation of G fixed, whence it follows easily that δ commutes with every left translation, so that δ lies in the center of the group of all proper automorphisms of $R(G)$.

On the other hand, by condition (1), we have $\mathrm{Hom}(G, C) \subset S$. Hence δ leaves the elements of $\mathrm{Hom}(G, C)$ fixed.

Let T be any finitely generated fully stable subalgebra of $R(G)$ containing S. Let M_T be a maximal fully reducible subgroup of A_T, so that A_T is the semidirect product $M_T \cdot U_T$. Accordingly, write $\delta_T = mu$, with $m \in M_T$ and $u \in U_T$. Since δ_T is in the center of M_T, m and u are then the semisimple and unipotent components of δ_T. In particular, this implies that u leaves the elements of $\mathrm{Hom}(G, C)$ fixed. We know from [1, p. 125] that the kernel of the representation of U on $C + \mathrm{Hom}(G, C)$ is precisely N. Hence we conclude that $u \in N_T$. Since δ_T leaves the elements of S fixed, so does u. Since S separates the elements of G, it follows that $u = 1$. Hence δ_T is semisimple, and we have shown that every proper automorphism leaving the elements of S fixed is a semisimple element of the center of the group of all proper automorphisms of $R(G)$.

Let F denote the group of all semisimple elements of the center of the group of all proper automorphisms of $R(G)$. What we have just proved is that the fixer S^* of S in the group of the proper automorphisms of $R(G)$ is contained in F. Hence $R(G)^F \subset R(G)^{S^*}$. Since S is fully stable and finitely generated, it follows from the standard results on algebraic linear groups and proper automorphisms that $R(G)^{S^*} = S$. Thus we have $R(G)^F \subset S$.

Now recall from [1, Ths. 5.1, 5.2] that, as a group of proper automorphisms of $R(G)$, GU has the form $M \cdot U$, where M is a fully reducible group of proper automorphisms, and that $R(G)^V(R(G)^M)'$ is then a normal basic subalgebra of $R(G)$. Now, if T is as above, then A_T is evidently the semidirect product of the algebraic hull of the image of M in A_T and U_T. Thus the algebraic hull of the image of M in A_T is a maximal fully reducible subgroup of A_T. It must therefore contain the image F_T of F in A_T. Hence we conclude that $R(G)^M \subset R(G)^F$.

Now we have $R(G)^M \subset S$, whence also $(R(G)^M)' \subset S$. Since $R(G)^V \subset S$, we finally conclude that S contains the normal basic subalgebra $R(G)^V(R(G)^M)'$, so that S is indeed a regular subalgebra of $R(G)$. This completes the proof of Theorem 4.1.

Theorem 4.1 can be strengthened considerably in the case where the given group G is solvable. In order to see this, let us first examine the general situation from a slightly different point of view. Let S be a finitely generated fully stable subalgebra of $R(G)$ that separates the points of G and contains $\mathrm{Hom}(G, C)$. Let K be a nucleus of G, and write $G = H \cdot K$, where H is a reductive closed analytic subgroup of G. Then H_S is contained in a maximal fully reducible subgroup M of the group A_S of all proper automorphisms of S. As before, we have $A_S = M \cdot U_S$ and $M \cap G_S = H_S$.

Now $\dim(A_S) = \dim(M) + \dim(U_S)$, and we know from [1, Cor. 5.1] that $\dim(U) = \dim(K)$. Since $\dim(U_S) \leqq \dim(U)$ and $\dim(K_S) = \dim(K)$, we have $\dim(A_S) \leqq \dim(M) + \dim(K_S)$. Since M and K_S are analytic subgroups of A_S, since $M \cap K_S = (1)$, and since A_S is connected, it follows that A_S is the semidirect product $M \cdot K_S$.

Now let us suppose that G is solvable. Then A_S must also be solvable, which implies that M is abelian. Since M is a fully reducible connected algebraic linear group, we know from [2, Lemma 2.2] that M is of the form $(C^*)^m$. The same holds for the connected algebraic subgroup H_S of M. Let P be the maximum compact subgroup of H_S, and let L be the maximum compact subgroup of M. Then P and L are toroidal groups and $P \subset L$. It follows immediately from the character theory that $L = P \times P_1$, where P_1 is again a toroidal group. Now M is the universal complexification of L and H_S is the universal complexification of P. Hence it is clear that M is the direct product $M_1 \times H_S$, where M_1 is the universal complexification of P_1.

In particular, our result shows that the canonical map $G \rightarrow A_S$ is a *regular* algebraic hull. By Theorem 4.1, S is therefore a *regular* subalgebra of $R(G)$. We may state these facts as follows.

THEOREM 4.2. *Let G be a solvable analytic group. Then every finitely* *generated fully stable subalgebra of* $R(G)$ *that separates the points of* G *and contains* $\mathrm{Hom}(G, C)$ *is a regular subalgebra of* $R(G)$, *i.e., contains a normal basic subalgebra of* $R(G)$. *Equivalently, every algebraic hull* $\rho\colon G \to L$ *such that every analytic homomorphism of* G *into* C *is the composite of* ρ *with a rational homomorphism of* L *into* C *is a regular algebraic hull.*

We remark that the example given at the end of [1, Section 3] shows that Theorem 4.2 does not hold if the solvability assumption is dropped. In fact, it is easily seen that the smallest fully stable algebra containing the algebra B of that example is not regular.

5. Nuclear subalgebras and nuclear algebraic hulls. Let K be any nucleus of G. It is clear from [1, Th. 4.1] that there is a unique regular subalgebra $R(G,K)$ of $R(G)$ such that, if B is any normal basic subalgebra of $R(G)$ whose associated nucleus is K, $R(G,K)$ coincides with the smallest fully stable subalgebra of $R(G)$ that contains B. We shall call $R(G,K)$ the *nuclear subalgebra* of $R(G)$ that is associated with K. We shall denote by $Q(K)$ the subgroup of Q that is generated by those elements q of Q whose restrictions q_K to K occur as components in the semisimple representation associated with the adjoint representation of K.

THEOREM 5.1. *Let* G *be a complex analytic group whose points are separated by* $R(G)$, *and let* K *be a nucleus of* G. *Let* B *be any normal basic subalgebra of* $R(G)$ *whose associated nucleus is* K. *Then* $R(G,K) = B[Q(K)]$.

Proof. By Lemma 3.1, $R(G,K) = B[Q_1]$, where $Q_1 = R(G,K) \cap Q$. What remains to be proved is only that $Q_1 = Q(K)$. By Lemma 3.2, $R(G,K)$ contains the representative functions associated with the adjoint representation of G. Now let q be an element of Q such that q_K occurs as a component of the semisimple representation associated with the adjoint representation of K. This representation is a component of the representation of K that is induced by the semisimple representation associated with the adjoint representation of G. Hence we conclude that $q_K \in R(G,K)_K$. We have $R(G,K)_K = B_K[(Q_1)_K]$, and the elements of Q_K are free over B_K. Hence it follows that $q_K \in (Q_1)_K$, which implies that $q \in Q_1$. Thus we have shown that $Q(K) \subset R(G,K)$.

In order to complete our proof, it will therefore suffice to show that $B[Q(K)]$ is fully stable. We know that it is stable under the left translations. In order to show that it is stable under the right translations, it suffices to show that, for every element g of G, $B \cdot g \subset B[Q(K)]$.

Let S be any finite-dimensional left K-stable subspace of B. By Corollary 2.1, S is the sum of K-submodules S_δ corresponding to equivalence classes δ, mod $Q(K)_K$, of analytic homomorphisms $q_K\colon K \to C^*$, where q is an element of Q. If $S_\delta \neq (0)$ there is a non-zero element s in S and an element q_K in δ such that $x \cdot s = q(x)s$ for every element x of K. This means that $q^{-1}s \in R(G)^K$, so that $q^{-1}s \in B$. Thus $s \in B \cap Bq$. But $B \cap Bq = (0)$, unless $q = 1$. Hence δ is necessarily the equivalence class containing 1, i.e., $\delta \subset Q(K)_K$, and $S = S_\delta$.

For every element q of Q, let ϕ_q denote the projection $R(G) \to Bq$ that corresponds to the decomposition $R(G) = \sum_{q \in Q} Bq$. If ρ_g denotes the right translation on $R(G)$ effected by the element g of G then $\phi_q \circ \rho_g$ is evidently a left G-module endomorphism of $R(G)$. In particular, $\phi_q(S \cdot g)$ is a K-homomorphic image of S. Hence it follows from our above result that every non-zero $\phi_q(S \cdot g)$ has only a single associated class of homomorphisms $K \to C^*$, and that this class is contained in $Q(K)_K$. On the other hand, $\phi_q(S \cdot g)$ is evidently of the form Tq, where T is a K-submodule of B. If τ is the class of homomorphisms $K \to C^*$ that is associated with T then the class of homomorphisms associated with Tq is evidently τq_K. Now we know that both τ and τq_K are contained in $Q(K)_K$. Hence we must have $q_K \in Q(K)_K$, whence $q \in Q(K)$. We have shown that if $\phi_q(S \cdot q) \neq (0)$ then $q \in Q(K)$. This gives $S \cdot g \subset B[Q(K)]$, whence $B \cdot g \subset B[Q(K)]$, which completes the proof that $B[Q(K)]$ is stable under the right translations.

It remains to show that $B[Q(K)]$ is stable under the involution $f \to f'$. Since $B[Q(K)]$ is two-sidedly stable, this will follow as soon as we have shown that, whenever $B[Q(K)]$ contains the representative functions associated with some representation ρ, it also contains the representative functions associated with the representation ρ^* dual to ρ. Now the representative functions associated with ρ^* belong to the algebra generated by the representative functions associated with ρ and the reciprocal of the function $d \circ \rho$, where d is the determinant function associated with ρ. We have a semidirect product decomposition $G = H \cdot K$. Let p be the element of $\mathrm{Hom}(G, C^*)$ defined by $p(hk) = d(\rho(h))$, where $h \in H$ and $k \in K$. Then $p \in R(G)^K$, and hence also $p^{-1} \in R(G)^K \subset B$. Now $p^{-1}(d \circ \rho) \in \mathrm{Hom}(G, C^*)$ and its restriction to H is the constant 1. Hence $p^{-1}(d \circ \rho) \in Q$. On the other hand, since $p^{-1} \in B$ and $d \circ \rho \in B[Q(K)]$, by our assumption on ρ, we have $p^{-1}(d \circ \rho) \in B[Q(K)]$. Hence $p^{-1}(d \circ \rho) \in Q(K)$. Thus the reciprocal of $d \circ \rho$ belongs to $p^{-1}Q(K) \subset B[Q(K)]$. This completes our proof of Theorem 5.1.

We write $A(G, K)$ for $A_{R(G,K)}$ and call the canonical map $\gamma\colon G \to A(G, K)$ the *nuclear algebraic hull* associated with the nucleus K. If $\rho\colon G \to L$ and

$\sigma: G \to M$ are any algebraic hulls, we shall say that ρ *dominates* σ if there is a rational homomorphism $\beta: L \to M$ such that $\beta \circ \rho = \sigma$. Then β is necessarily surjective, and β is unique. As an immediate consequence of Theorems 4.1 and 3.1 we have the following result.

THEOREM 5.2. *An algebraic hull $\rho: G \to L$ is regular if and only if it dominates a nuclear algebraic hull $\gamma: G \to A(G,K)$ associated with a nuceus K of G.*

We wish to characterize the nuclear algebraic hulls $\gamma: G \to A(G,K)$ by a certain universal property. For any algebraic linear group P, let P^* denote the factor group of P mod its maximum normal unipotent subgroup, and let δ^P be the canonical rational epimorphism $P \to P^*$. Let α denote the adjoint representation of K, and let $[\alpha(K)]$ be the algebraic group hull of $\alpha(K)$ in the group of all automorphisms of the Lie algebra of K. Now let $\phi: G \to L$ be an analytic homomorphism, where L is an algebraic linear group and $\phi(G)$ is algebraically dense in L. We shall say that ϕ is *K-special* if there is a rational homomorphism $\phi^+: [\alpha(K)]^* \to L^*$ such that $\phi^+ \circ \delta^{[\alpha(K)]} \circ \alpha = \delta^L \circ \phi_K$.

THEOREM 5.3. *In the notation introduced above, the canonical homomorphism $\gamma: G \to A(G,K)$ is K-special. If $\phi: G \to L$ is any K-special homomorphism there exists one and only one rational homomorphism ϕ^γ: $A(G,K) \to L$ such that $\phi = \phi^\gamma \circ \gamma$.*

In order to prove this result, we require some discussion of semisimple representative functions. Let ϕ be a complex analytic representation of G, and let L be the algebraic group hull of $\phi(G)$ in the requisite full linear group. Let $S(\phi)$ denote the smallest fully stable subalgebra of $R(G)$ that contains the representative functions associated with ϕ. Let $P(L)$ denote the algebra of all rational representative functions on L. It is seen immediately from the definitions that the map $f \to f \circ \phi$ is an algebra isomorphism of $P(L)$ onto $S(\phi)$.

Let U denote the maximum unipotent normal subgroup of L, i.e., the kernel of δ^L. Then the U-fixed subalgebra $P(L)^U$ of $P(L)$ is isomorphic, by the cohomomorphism of δ^L, with $P(L^*)$. Since L^* is a fully reducible algebraic linear group, it follows that $P(L)^U$ consists entirely of semisimple representative functions. Hence $P(L)^U \circ \phi$ is contained in the algebra $S(\phi)_s$ of the semisimple elements of $S(\phi)$. Conversely, if f is an element of $P(L)$ such that $f \circ \phi$ is a semisimple representative function on G it follows from the fact that $\phi(G)$ is algebraically dense in L that f is a semisimple representative function on L. This implies that the representation of U on the space of the left translates of f is semisimple and hence that $f \in P(L)^U$. Hence we have $S(\phi)_s = P(L)^U \circ \phi$.

If V is the representation space for ϕ then the members of a composition series for V as a G-module are stable also under L. Let V' be the direct sum of the factor modules of this composition series. Then, as a G-module, V' realizes the semisimple representation ϕ' associated with ϕ. On the other hand, the representation of L on V' has precisely U for its kernel. The image of L under this representation may be identified with L^*, and if we apply the above consideration to ϕ' and L^* we find that $S(\phi')$ is isomorphic with $P(L^*)$.

We may summarize these facts by writing $S(\phi)_s = P(L^*) \circ \delta^L \circ \phi = S(\phi')$. The following lemma will be convenient for the proof of Theorem 5.3.

LEMMA 5.1. *A complex analytic representation ϕ of G is K-special if and only if the restrictions to K of the semisimple representative functions associated with ϕ belong to $S(\alpha)$, where α denotes the adjoint representation of K.*

Proof. Suppose first that $\phi\colon G \to L$ is K-special. Let f be a semisimple representative function associated with ϕ. From the above, we know that $f = g \circ \delta^L \circ \phi$, with g in $P(L^*)$. Hence $f_K = g \circ \delta^L \circ \phi_K$. By assumption, we have $\delta^L \circ \phi_K = \phi^+ \circ \delta^{[\alpha(K)]} \circ \alpha$, where ϕ^+ is a rational homomorphism $[\alpha(K)]^* \to L^*$. Thus we have $f_K = (g \circ \phi^+ \circ \delta^{[\alpha(K)]}) \circ \alpha$. Since $\phi^+ \circ \delta^{[\alpha(K)]}$ is a rational homomorphism, we have $g \circ \phi^+ \circ \delta^{[\alpha(K)]} \in P([\alpha(K)])$, so that the last result gives $f_K \in S(\alpha)$.

Now suppose that the restriction to K of every semisimple representative function associated with ϕ belongs to $S(\alpha)$. From our above discussion, we know that $S(\delta^L \circ \phi) = S(\phi)_s$. Hence our assumption gives $S(\delta^L \circ \phi)_K \subset S(\alpha)_s$. The result just used for ϕ holds also for α, so that $S(\alpha)_s = S(\delta^{[\alpha(K)]} \circ \alpha)$. Thus we have $S(\delta^L \circ \phi_K) \subset S(\delta^{[\alpha(K)]} \circ \alpha)$. Let $x \in [\alpha(K)]^*$. Then we may regard x as a specialization of $P([\alpha(K)]^*)$. The map $f \to f \circ \delta^{[\alpha(K)]} \circ \alpha$ is an isomorphism of $P([\alpha(K)]^*)$ onto $S(\delta^{[\alpha(K)]} \circ \alpha)$ by means of which x is transported to a specialization x' of the last algebra. By restriction, x' induces a specialization x'' of $S(\delta^L \circ \phi_K)$. Via $\delta^L \circ \phi_K$, this yields a specialization x''' of $P(L^*)$ and thus an element $\phi^+(x)$ of L^*. It is clear from this construction that ϕ^+ is an everywhere defined rational map of $[\alpha(K)]^*$ into L^*.

On the other hand, since $S(\delta^L \circ \phi_K) \subset S(\delta^{[\alpha(K)]} \circ \alpha)$, the kernel of $\delta^L \circ \phi_K$ contains the kernel of $\delta^{[\alpha(K)]} \circ \alpha$, so that $\delta^L \circ \phi_K$ induces a homomorphism $\psi\colon \delta^{[\alpha(K)]}(\alpha(K)) \to L^*$. It is clear from the definitions that the restriction to $\delta^{[\alpha(K)]}(\alpha(K))$ of the above map ϕ^+ coincides with this homomorphism ψ. Since $\delta^{[\alpha(K)]}(\alpha(K))$ is algebraically dense in $[\alpha(K)]^*$, it follows that ϕ^+ is actually a homomorphism and that $\phi^{+} \circ \delta^{[\alpha(K)]} \circ \alpha = \delta^{L} \circ \phi_{K}$. Thus ϕ is K-special, and Lemma 5.1 is proved.

Now we are in a position to prove Theorem 5.3. We have $S(\gamma) = R(G,K)$, and $R(G,K)_s = (B[Q(K)])_s = B_s[Q(K)] = R(G)^K[Q(K)]$. Hence the restrictions to K of the semisimple representative functions associated with γ belong to $C[Q(K)_K]$. Since this last algebra is contained in $S(\alpha)$, it follows from Lemma 5.1 that γ is K-special.

Now let $\phi\colon G \rightarrow L$ be any K-special homomorphism. Let F denote the space of the representative functions associated with ϕ. We wish to show that $F \subset R(G,K)$. Clearly, F is contained in a sum $S_1q_1 + \cdots + S_nq_n$, where the S_i's are finite-dimensional two-sidedly translation stable subspaces of $R(G,K)$ and $q_1, \cdots, q_n$ are elements of Q no two of which belong to the same coset of $Q(K)$ in Q. Our sum is necessarily a direct sum. Let p_i be the projection of F to the summand S_iq_i. Since p_i commutes with the translations, we have $p_i(F) \subset F$. Hence F is actually equal to a sum $S_1q_1 + \cdots + S_nq_n$. Passing to the semisimple parts, we obtain

$$F_s = (S_1)_sq_1 + \cdots + (S_n)_sq_n.$$

Since ϕ is K-special, we have, from Lemma 5.1, $(F_s)_K \subset S(\alpha)_s$. But $S(\alpha)_s = S(\alpha') = C[Q(K)_K]$. Hence we conclude that $((S_i)_s)_K(q_i)_K \subset C[Q(K)_K]$. Since S_i is two-sidedly translation stable, it coincides with the space of all representative functions associated with the representation of G by left translations on S_i. Hence, if $S_i \neq (0)$, then also $(S_i)_s \neq (0)$. Since $(S_i)_s$ is translation stable, it follows that $((S_i)_s)_K \neq (0)$. On the other hand, since $S_i \subset R(G,K)$, we have $((S_i)_s)_K \subset C[Q(K)_K]$. Hence the inclusion we obtained above implies that $(q_i)_K \in Q(K)_K$, so that $q_i \in Q(K)$. This means that $F \subset R(G,K)$.

Hence we have $S(\phi) \subset R(G,K)$, and now we can produce the required rational homomorphism ϕ^{γ} by the process used in the proof of Lemma 5.1: let $x \in A(G,K)$, and let x' be the specialization of $S(\phi)$ obtained as the restriction of the specialization $f \rightarrow f(x)$ of $R(G,K)$. Via ϕ, this yields a specialization x'' of $P(L)$ and thus an element of L. Evidently, ϕ^{γ} is an everywhere defined rational map of $A(G,K)$ into L, and $\phi^{\gamma} \circ \gamma = \phi$. Since $\gamma(G)$ is algebraically dense in $A(G,K)$, it follows that ϕ^{γ} is actually a homomorphism and that it is uniquely determined as a rational homomorphism by the requirement $\phi^{\gamma} \circ \gamma = \phi$. The proof of Theorem 5.3 is therefore complete.

THEOREM 5.4. *Let K be a nucleus of the complex analytic group G.*

Let T denotes the maximum nilpotent normal analytic subgroup of K, and let U denote the terminal member of the descending central series of K. Finally, let U' denote the commutator subgroup of U. Then, if d is the rank of $Q(K)$, one has $\dim(K/T) \leqq d \leqq \dim(U/U')$.

Proof. Since the adjoint representation of K induces a unipotent representation of K on the Lie algebra of K/U, the non-trivial part of the associated semisimple representation is already obtained from the adjoint representation of K on the Lie algebra $\mathfrak{U}$ of U. Next we observe that if S is a subspace of $\mathfrak{U}$ such that $S + [\mathfrak{U}, \mathfrak{U}] = \mathfrak{U}$ then, since $\mathfrak{U}$ is nilpotent, we obtain $\mathfrak{U} = S_1 + \cdots + S_k$, where $S_1 = S$ and $S_{i+1} = [S, S_i]$. Hence we see that the semisimple K-module associated with the adjoint representation of K on $\mathfrak{U}$ is a homomorphic image of the direct sum of the i-th tensor powers $(i = 1, \cdots, k)$ of the semisimple K-module associated with the adjoint representation of K on $\mathfrak{U}/[\mathfrak{U}, \mathfrak{U}]$. Hence $Q(K)_K$ is already generated by the components of the semisimple representation associated with the adjoint representation of K on $\mathfrak{U}/[\mathfrak{U}, \mathfrak{U}]$, so that we must have $d \leqq \dim(\mathfrak{U}/[\mathfrak{U}, \mathfrak{U}]) = \dim(U/U')$.

On the other hand, we may view the semisimple representation of K that is associated with the adjoint representation of K as a semisimple representation σ of the vector group K/T. Since T is the maximum nilpotent normal analytic subgroup of K, the kernel of σ must be discrete. Hence $\sigma(K/T)$ is a group of the same dimension as K/T. Since $Q(K)_K$ is the group generated by the 1-dimensional components of σ, it is clear that σ may be written as a composite $\alpha \circ \beta$, where β is the analytic monomorphism of K/T into $(C^*)^d$ that is obtained from a free basis of $Q(K)$ and α is an analytic homomorphism of $(C^*)^d$ into the requisite full linear group. Hence the dimension of $\sigma(K/T)$ cannot exceed d, so that $\dim(K/T) \leqq d$. This completes the proof of Theorem 5.4.

Note that in the case where $T = K$, i.e., when K is nilpotent, we have necessarily $d = 0$. Hence Theorem 3.1 gives $\gamma(G) = A(G, K)$, in this case, and the above results contain our former result [1, Th. 8.3] that a nilpotent nucleus of G determines an algebraic structure of G.

The following simple example illustrates the variability of d within the limits given in Theorem 5.4. Let G be the group of all 4-tuples (x, a, b, c), where x ranges over the non-zero complex numbers and a, b, c range over all complex numbers, and where the multiplication is

$$(x, a, b, c)(x_1, a_1, b_1, c_1) = (xx_1, a + a_1, b \exp(a_1) + b_1, cx_1 + c_1).$$

Here, the radical N of the commutator subgroup of G consists of all elements of the form $(1,0,b,c)$, so that $\dim(N)=2$. Corresponding to an arbitrary complex number t, we define a nucleus K_t of G as the set of all elements of the form $(\exp(ta),a,b,c)$. For each t, the maximum nilpotent normal analytic subgroup T_t of K_t coincide with N, so that $\dim(K_t/T_t)=1$. In the notation of Theorem 5.4 (with the subscript t added), U_t consists of the elements $(1,0,b,c)$ if $t\neq 0$ (while U_0 consists of the elements $(1,0,b,0)$), and $U_t'=(1)$. Thus, for $t\neq 0$, $\dim(U_t/U_t')=2$. Now if t is a rational number the rank d_t of $Q(K_t)$ is equal to 1; if t is not rational then the rank d_t is equal to 2.

Moreover, if u and v are non-zero integers and $v>1$ then $R(G,K_u)$ contains $R(G,K_{uv})$ properly, showing that, contrary to what one might expect in view of Theorem 5.2, the nuclear algebraic hulls can be quite far from minimal, in the sense of domination.

Appendix. We take this opportunity to correct a confusion in a proof contained in [1]. In [1], p. 125, lines 4 and 3 from the bottom, replace *linearly independent* with *non-zero*. On line 2 from the bottom, insert *distinct* before *elements*. On p. 126, replace lines 3, 4, 5, 6 with the following.

Since $G\subset UM$ and $0\neq f_i'\in R^M$, there is an element u in U such that

$u(f_1')(1)\neq 0$. Put $g=\sum_{i=1}^{n}u(f_i')(1)q_i^{-1}$. Then $m(g)(1)=0$ for every

element m of M. But this gives $g=0$, contradicting the linear independence of $(q_1^{-1},\cdots,q_n^{-1})$.

REFERENCES.

[1] G. Hochschild and G. D. Mostow, "On the algebra of representative functions of an analytic group," *American Journal of Mathematics*, vol. 83 (1961), pp. 111-136.

[2] ———, "Representations and representative functions of Lie groups, III," *Annals of Mathematics*, vol. 70 (1959), pp. 85-100.

ERRATA.

Errata to A. H. Stone, "Hereditarily compact spaces," *American Journal of Mathematics,* vol. 82 (1960), pp. 900-916.

On p. 901, property (B_3) should read: "Every countable *open* covering . . ."; and (D_2) should read: "Every countably infinite *locally finite* open covering . . .". The italic words were inadvertently omitted.